Karlsruhe Institute of Technology

Please checkout the accepted manuscript (no paywall):
https://edouardfouche.com/publications/

DEAL: Data-Efficient Active Learning for regression under drift

A First Method Aimed at the Gap of Regression Active Learning with Drift

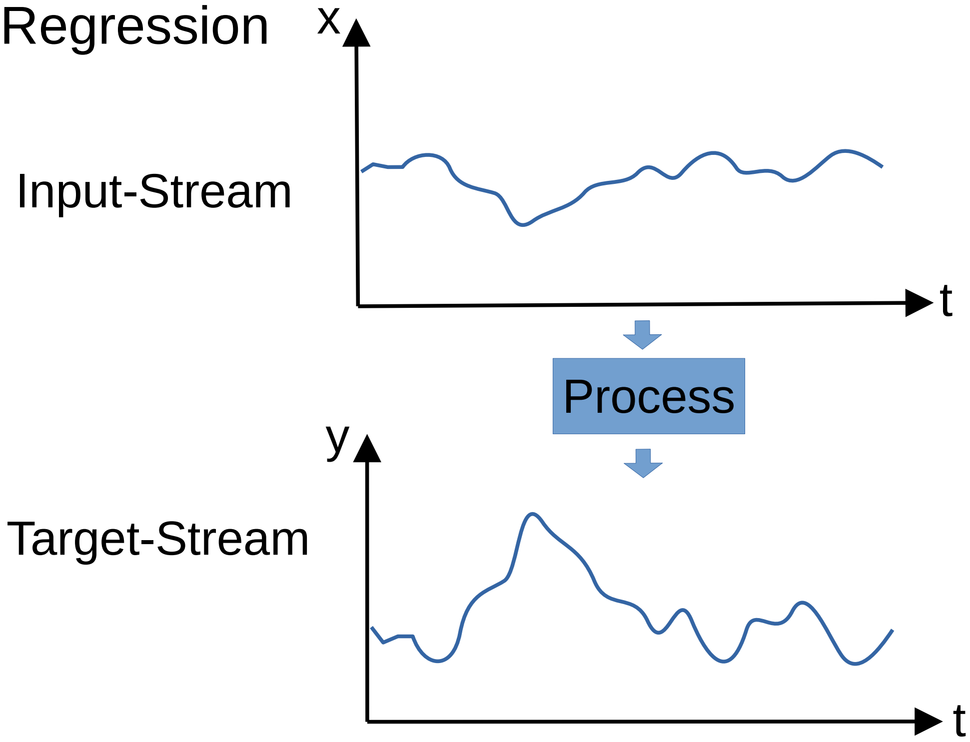
By Béla H. Böhnke, Edouard Fouché, and Klemens Böhm

Overview

- Problem Statement
  - Stream Regression
  - Active Learning for Regression
  - Active Learning for Regression under Drift
- Related Work
- Our Method
- Evaluation
- Outlook

Problem Statement

- Stream Regression



Problem Statement

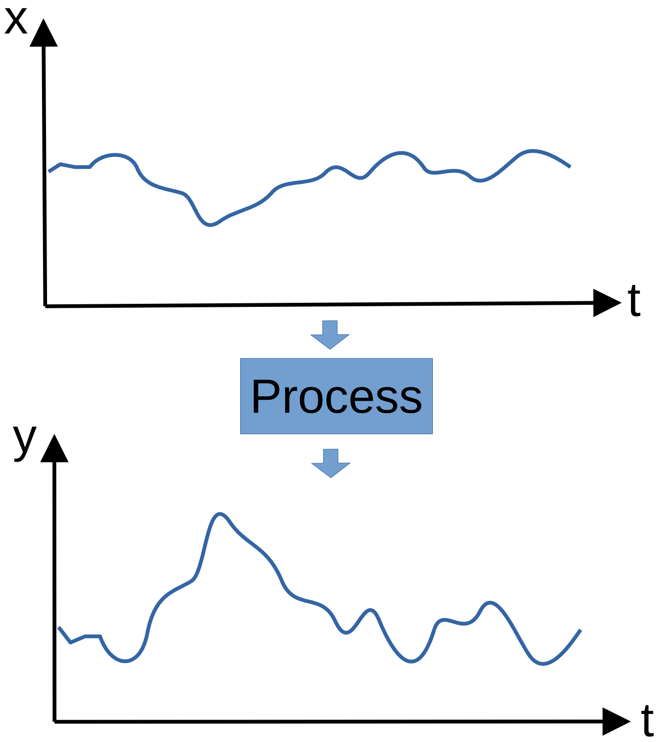
Karlsruhe Institute of Technology

- Stream Regression

Example:

Industrial process control

Input-Stream



Target-Stream

Problem Statement

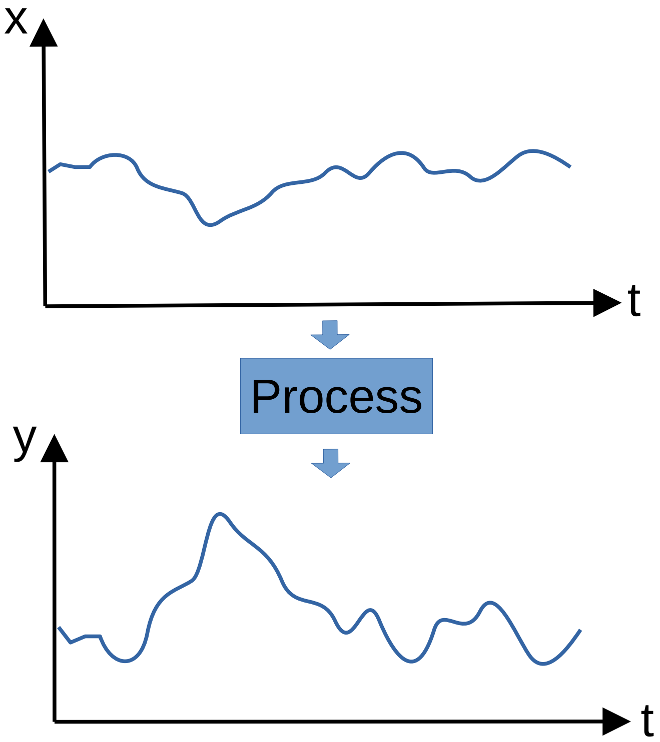
- Stream Regression

Example:
Industrial process control

Input-Stream

Process parameters like pressure, temperature and materials

Target-Stream



Problem Statement

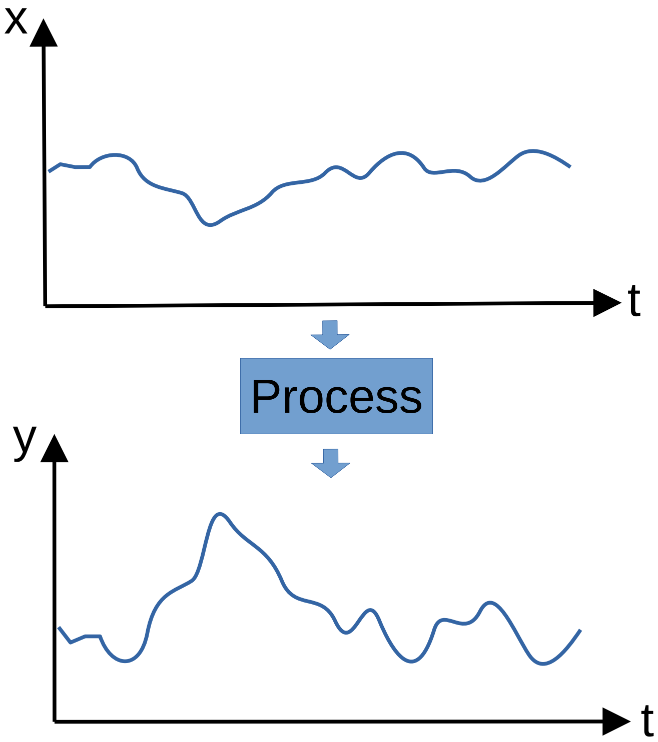
- Stream Regression

Example:

Industrial process control

Input-Stream

Process parameters like pressure, temperature and materials

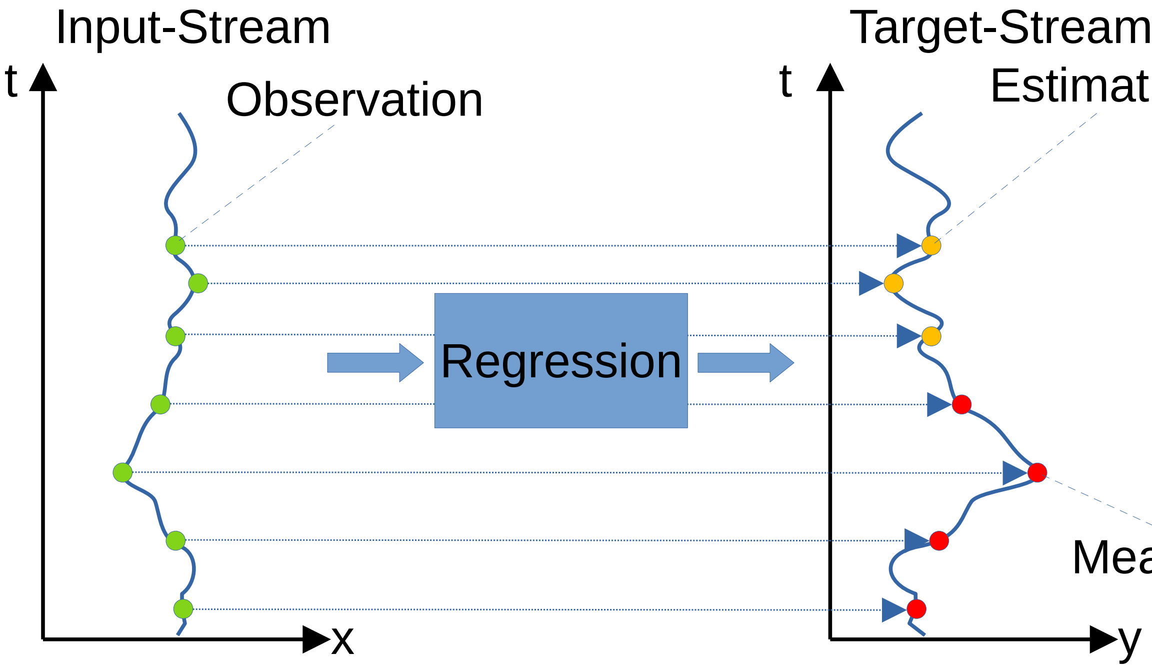


Target-Stream

Product quality

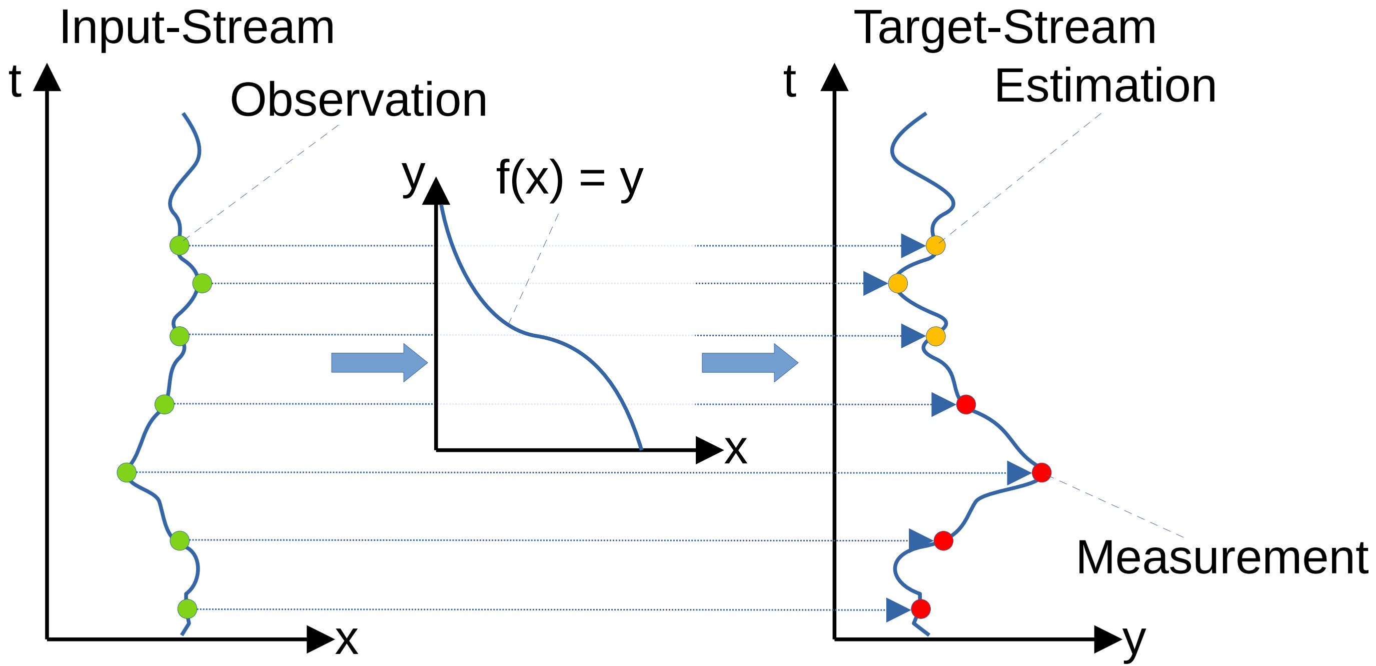
Problem Statement

- Stream Regression



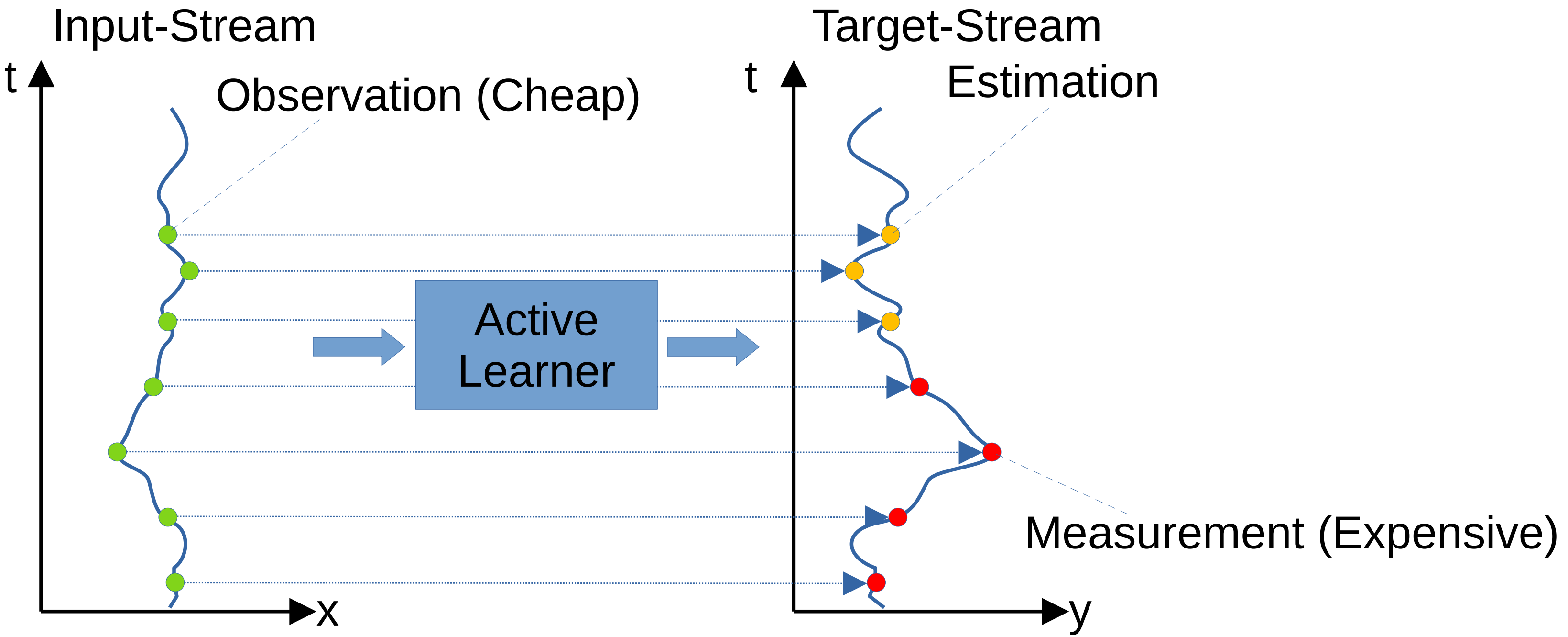
Problem Statement

- Stream Regression



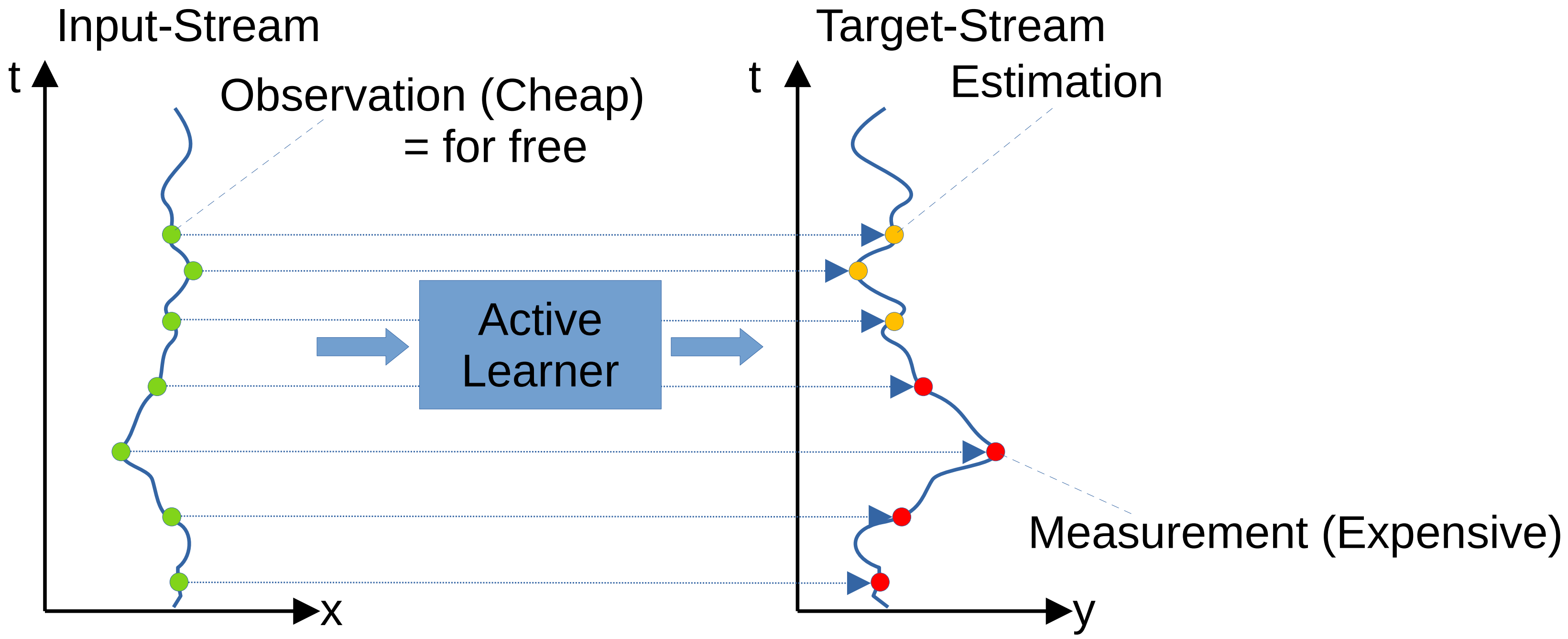
Problem Statement

- Active Learning for Regression



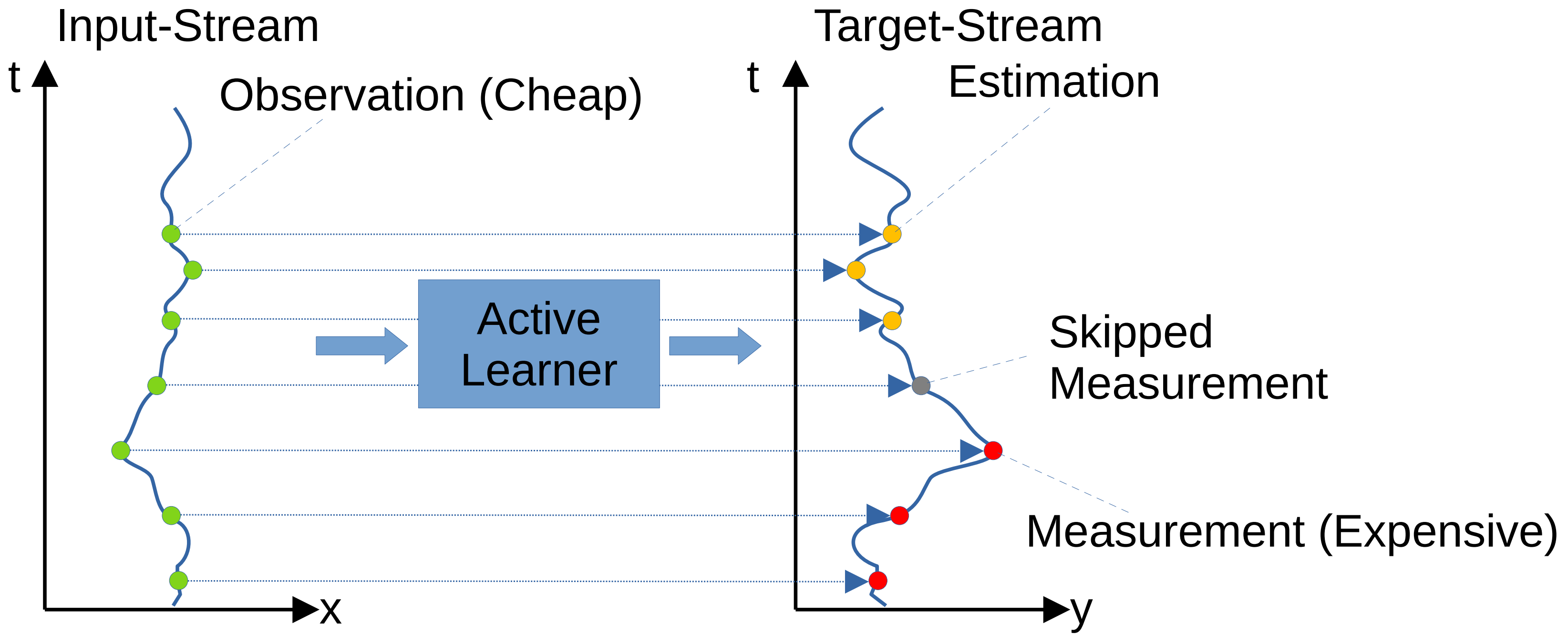
Problem Statement

- Active Learning for Regression



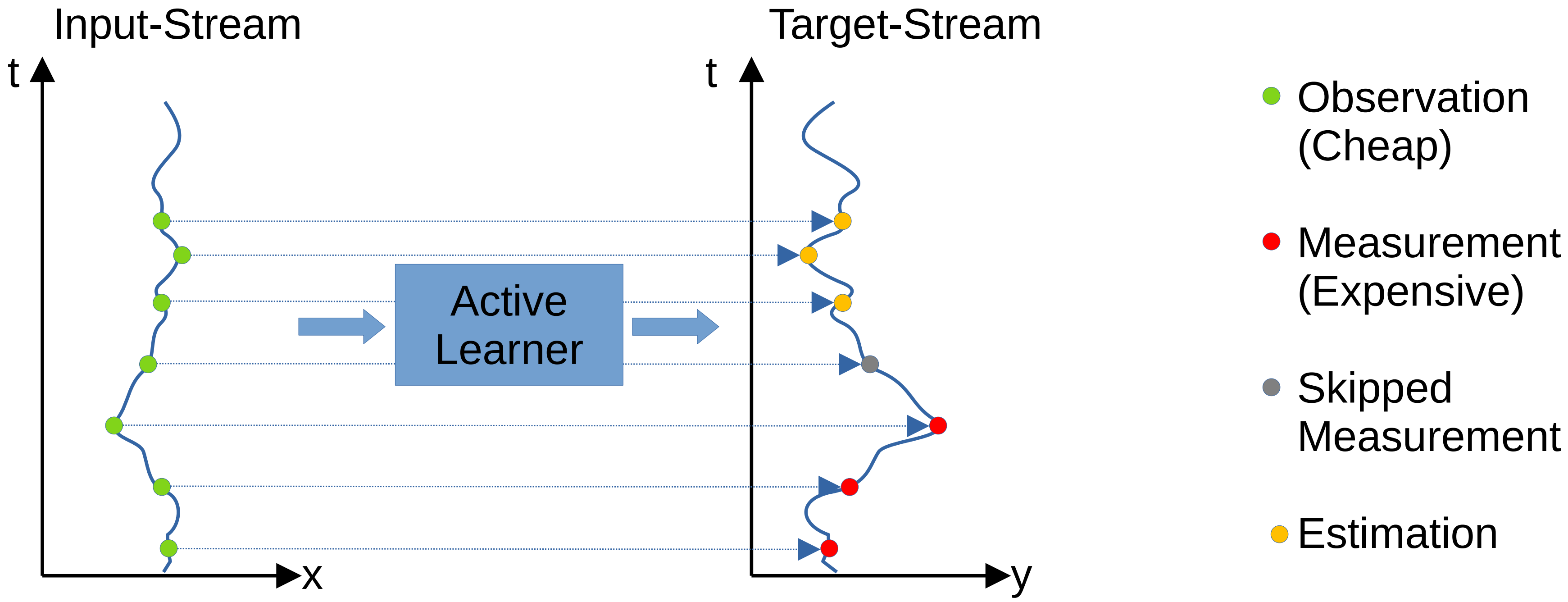
Problem Statement

- Active Learning for Regression



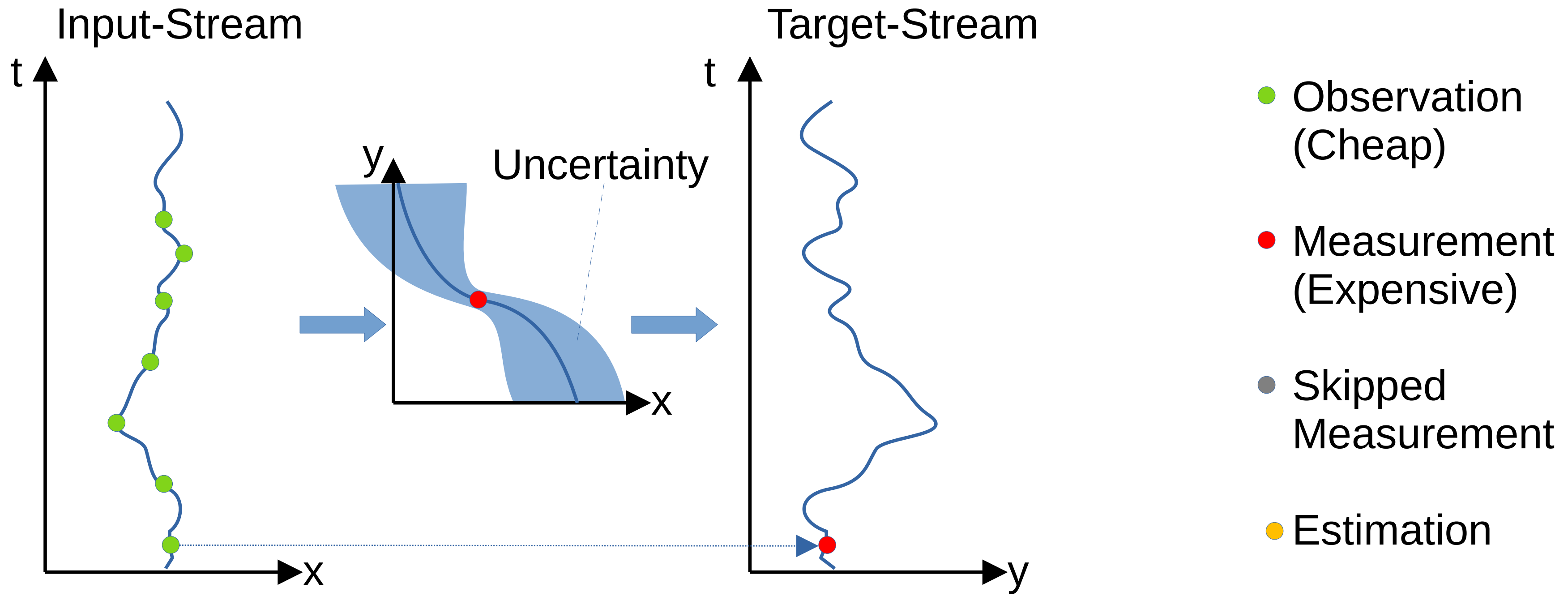
Problem Statement

- Active Learning for Regression



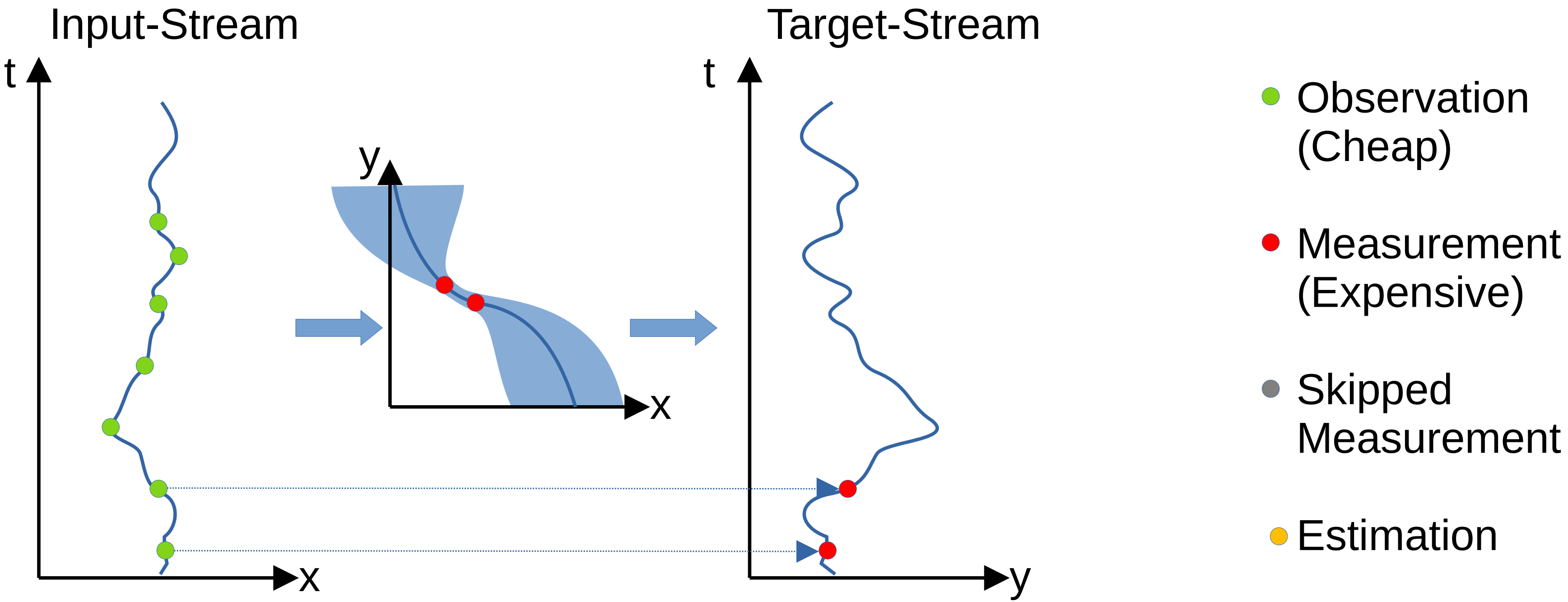
Problem Statement

- Active Learning for Regression



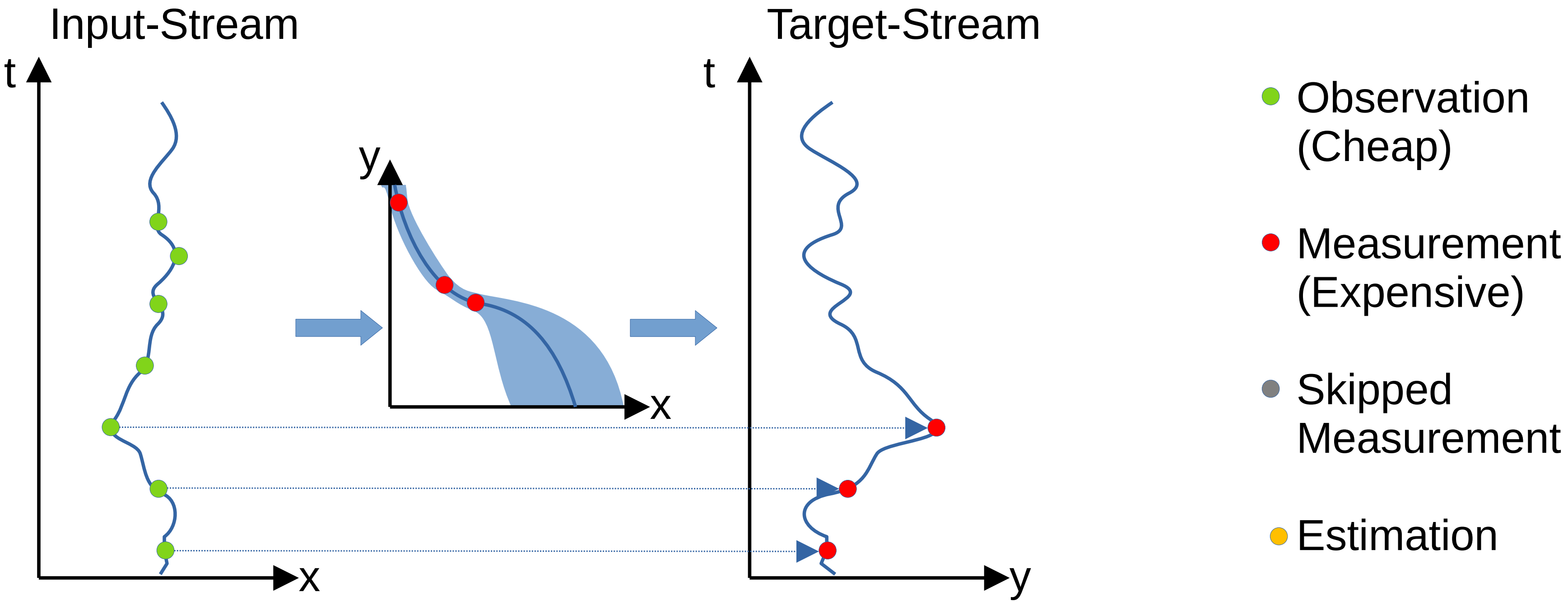
Problem Statement

- Active Learning for Regression



Problem Statement

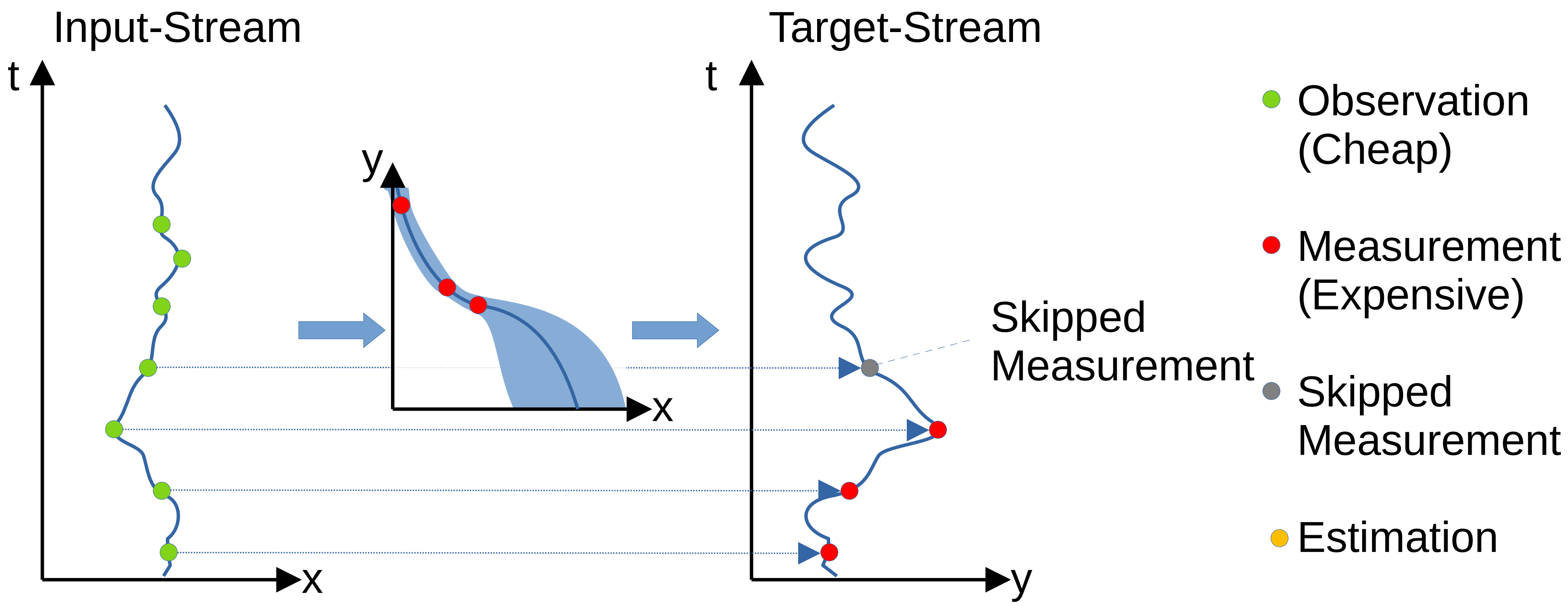
- Active Learning for Regression



Problem Statement

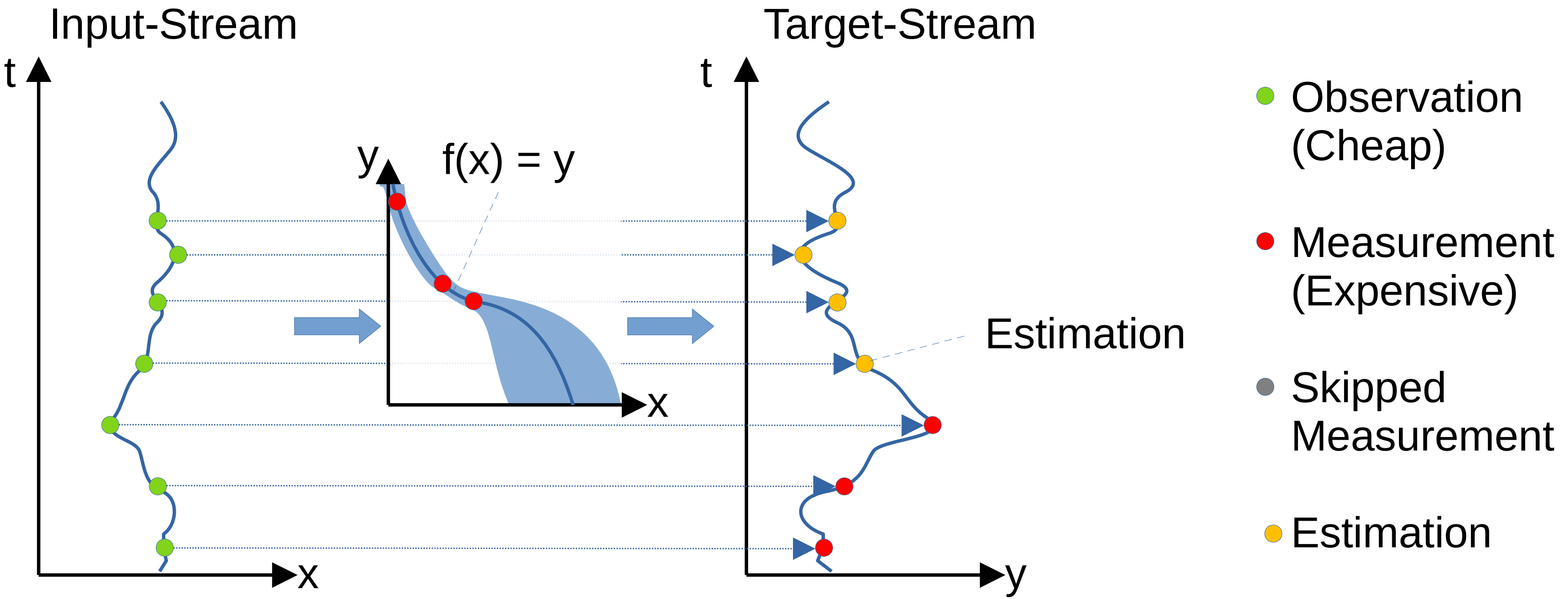
Karlsruhe Institute of Technology

- Active Learning for Regression



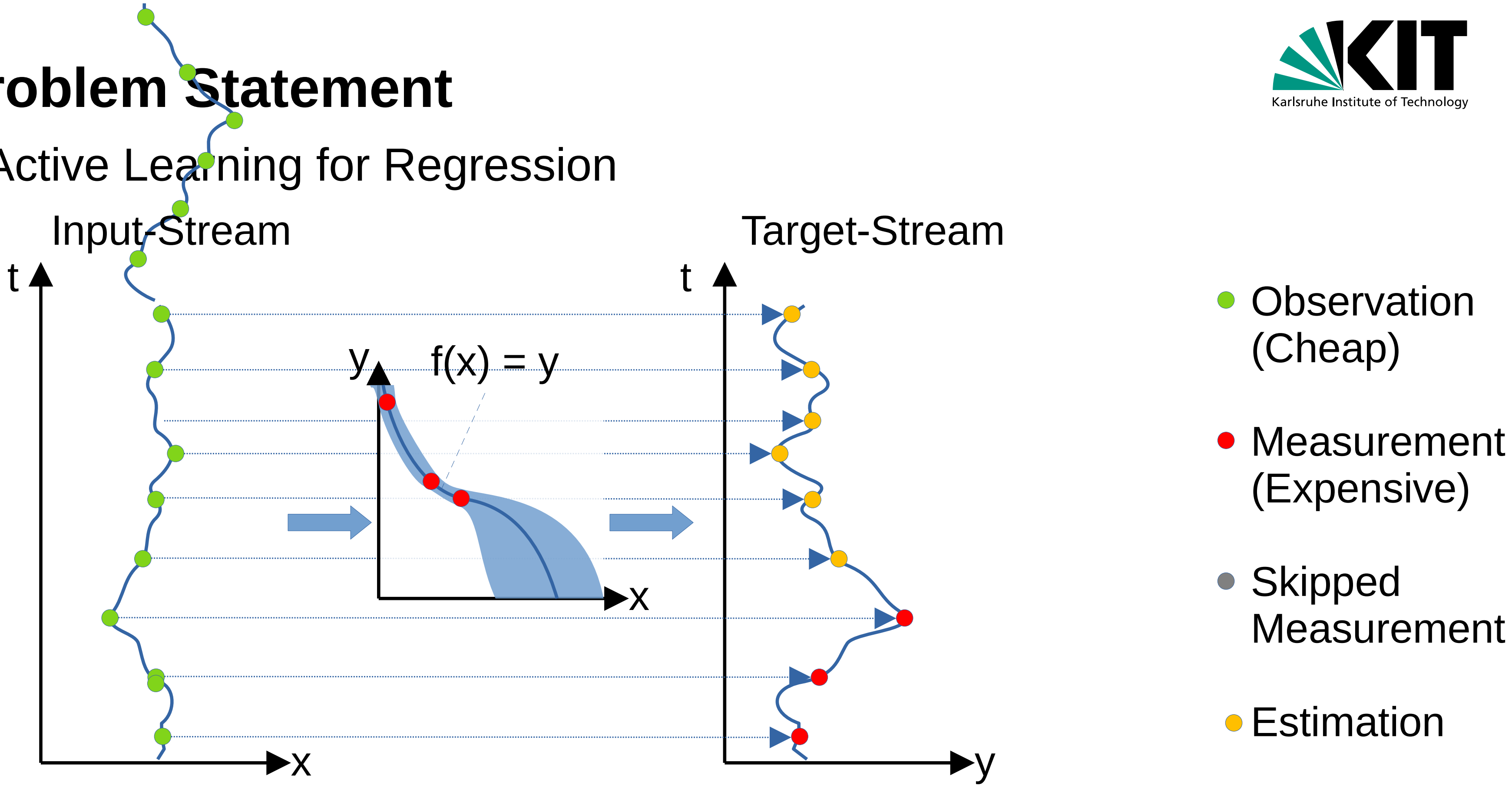
Problem Statement

- Active Learning for Regression



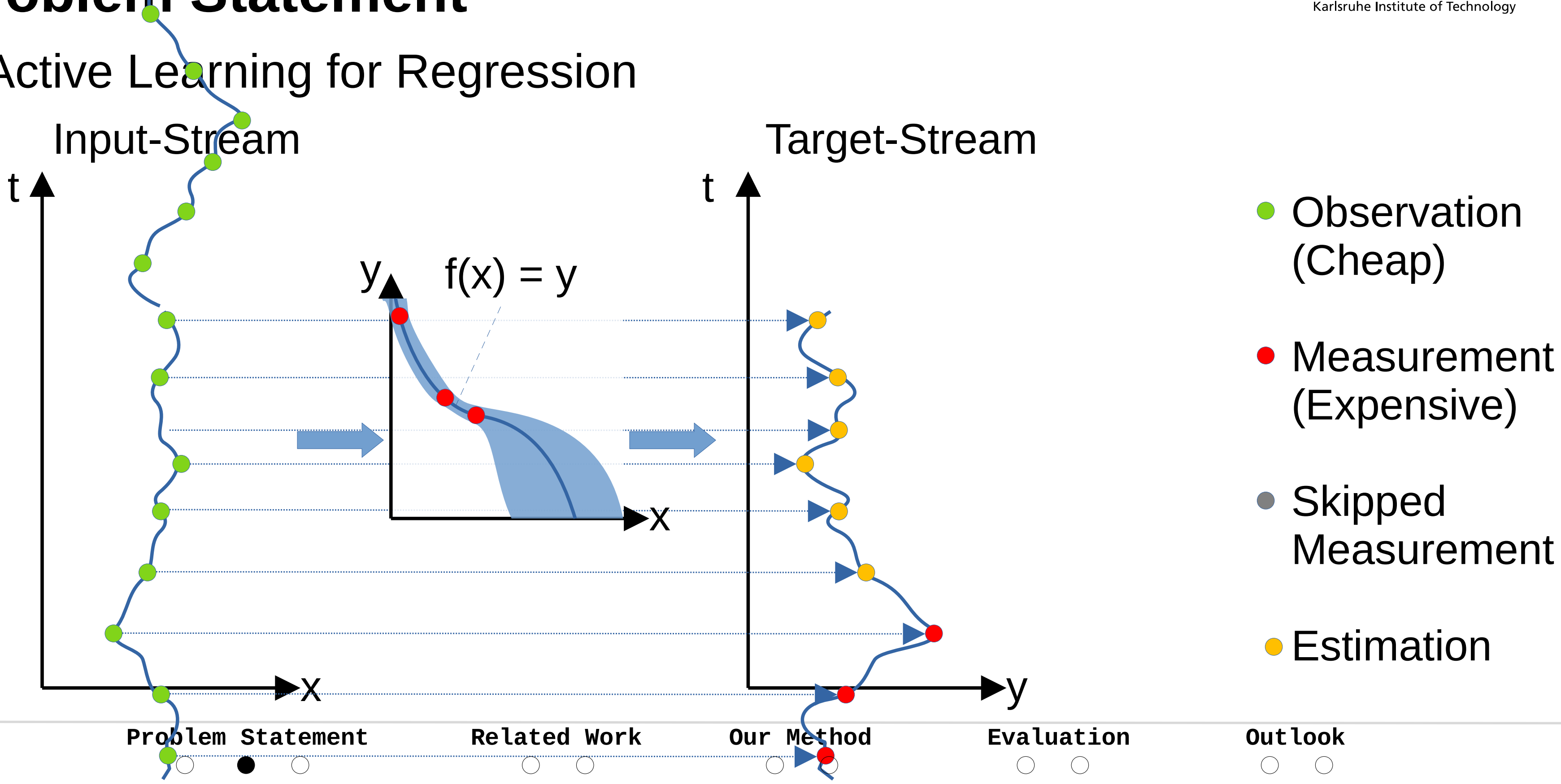
Problem Statement

- Active Learning for Regression



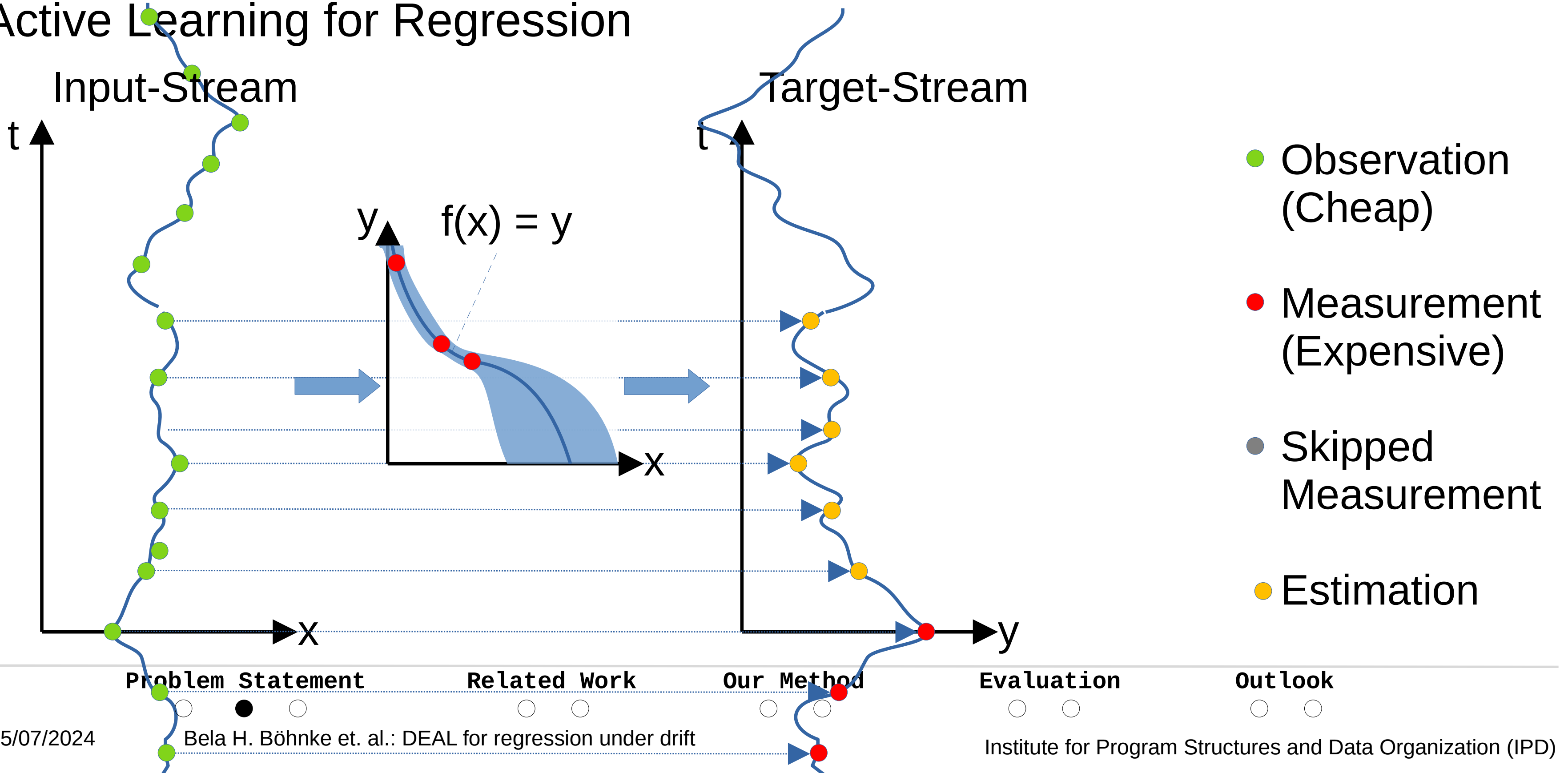
Problem Statement

- Active Learning for Regression



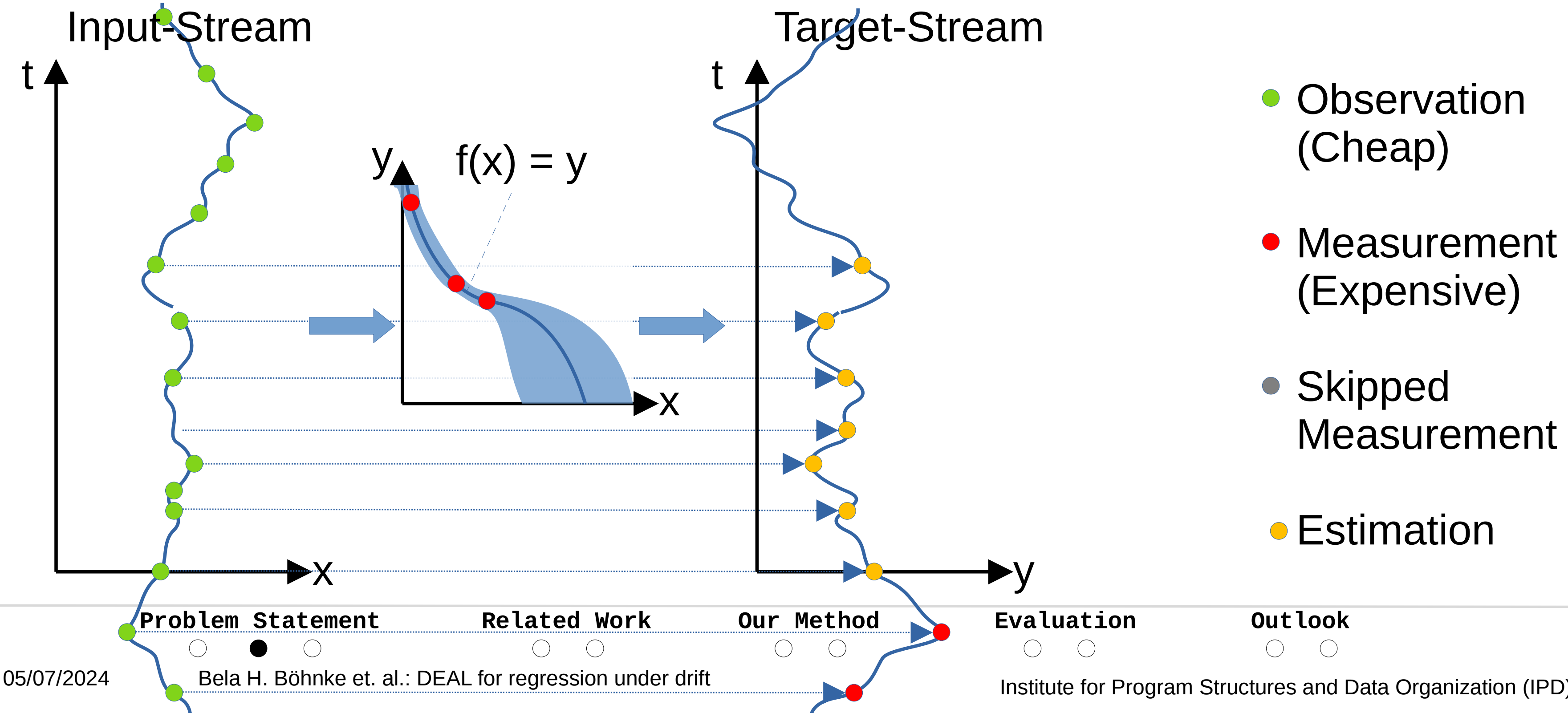
Problem Statement

- Active Learning for Regression



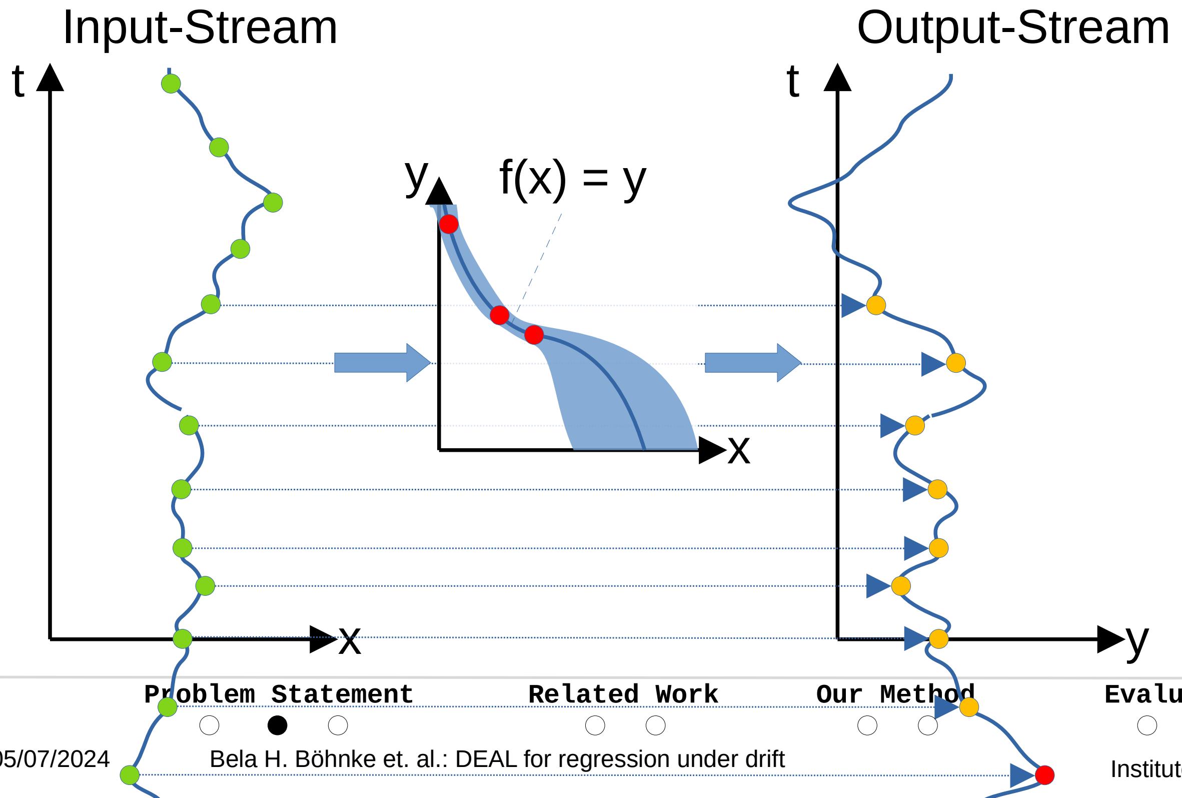
Problem Statement

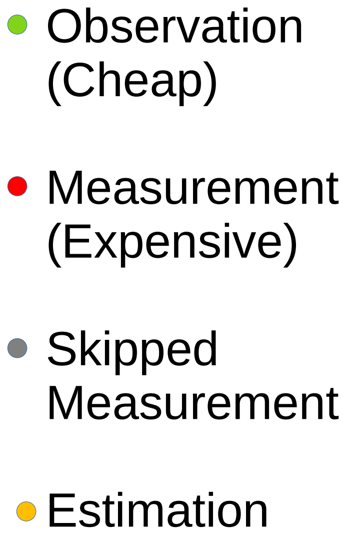
- Active Learning for Regression



Problem Statement

- Active Learning for Regression

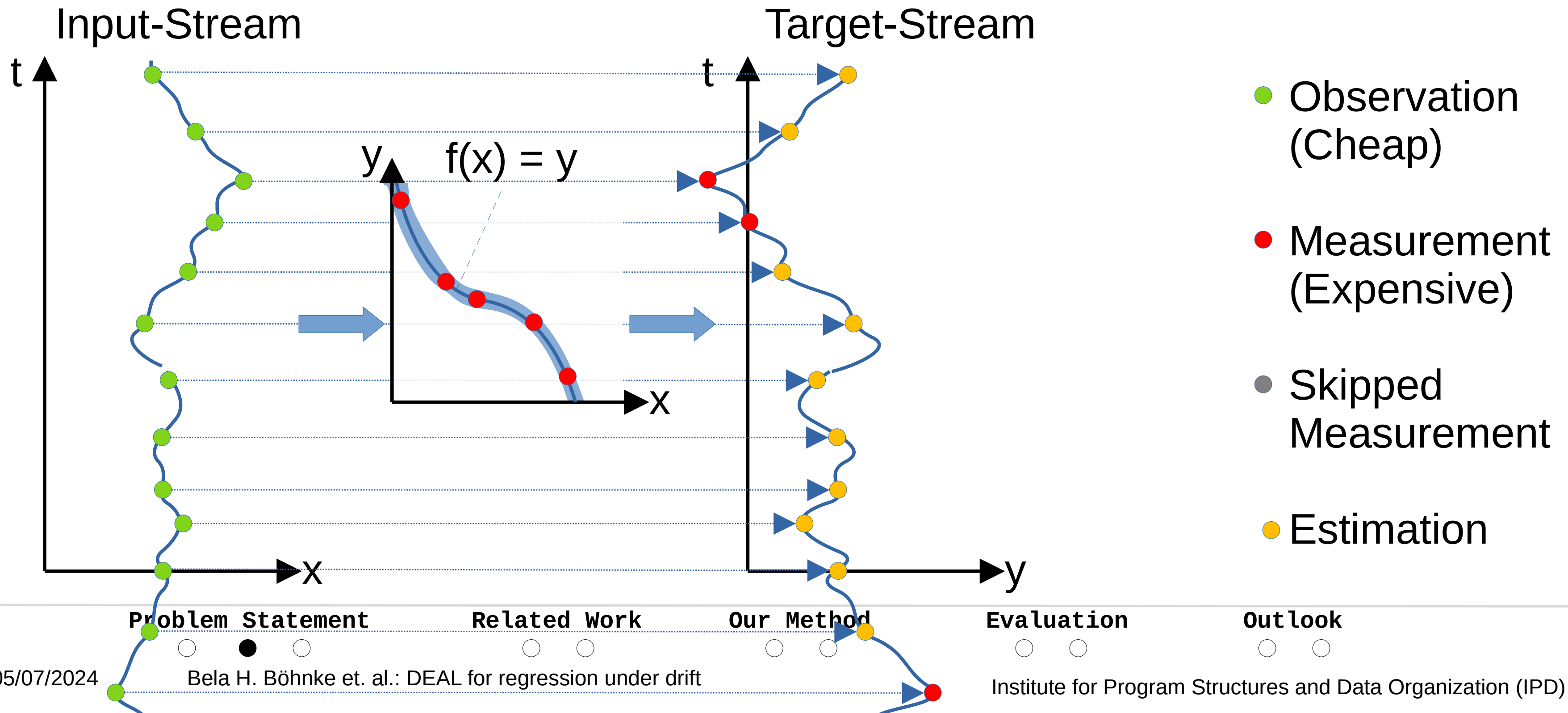




Problem Statement

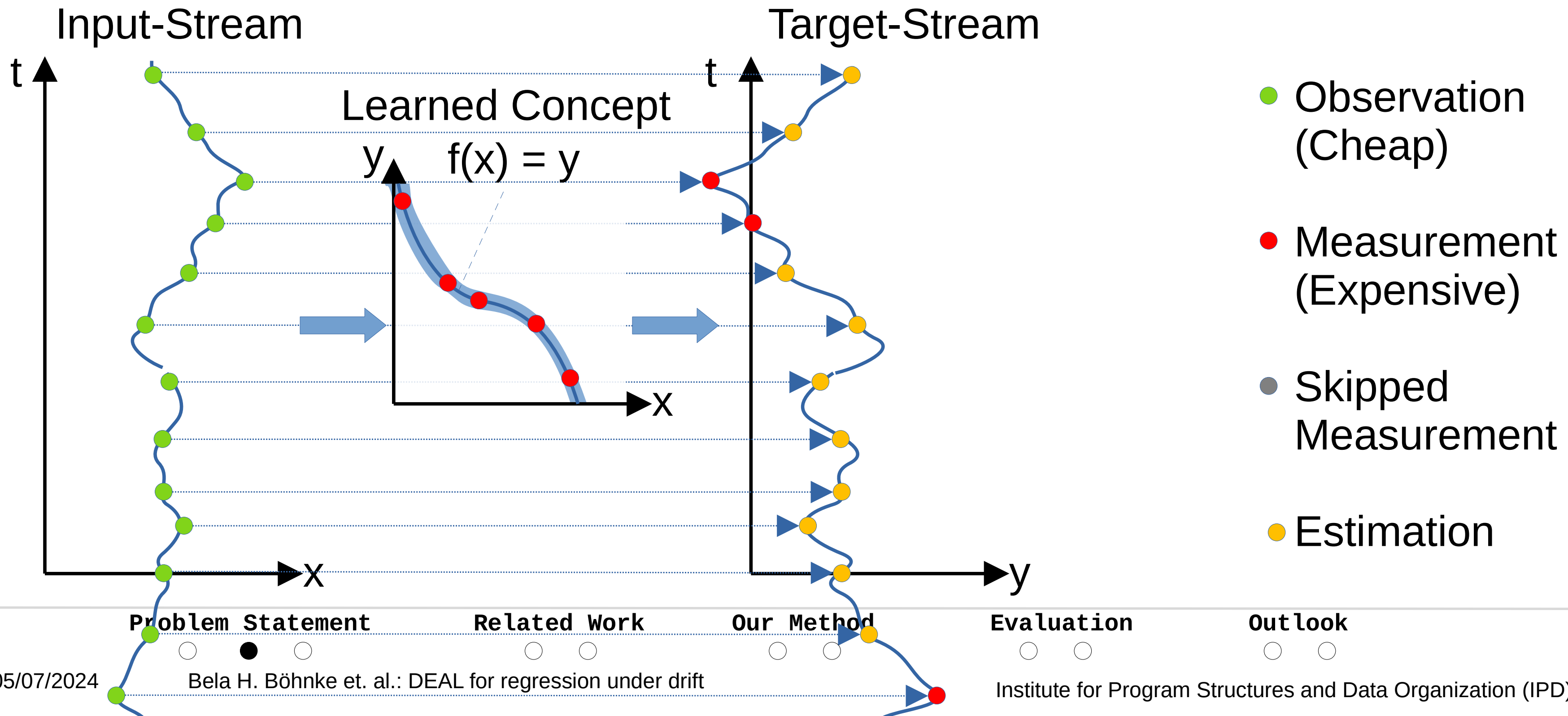
Karlsruhe Institute of Technology

- Active Learning for Regression



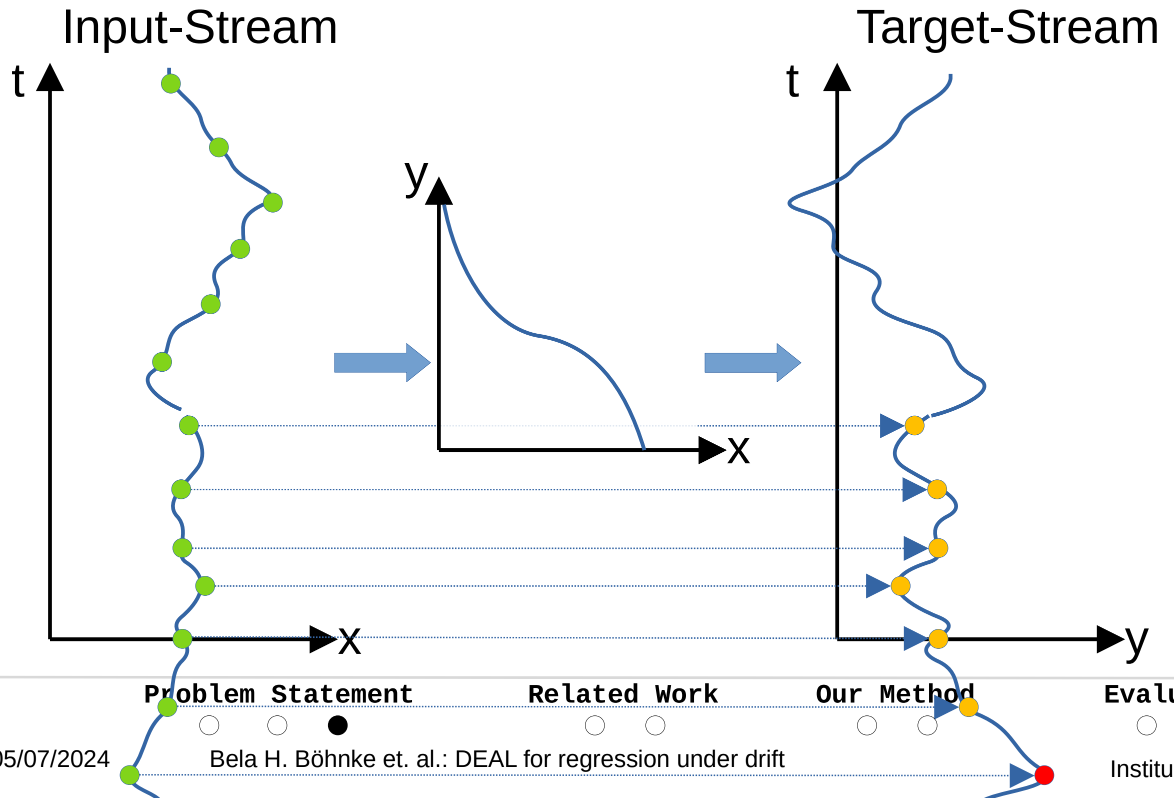
Problem Statement

- Active Learning for Regression



Problem Statement

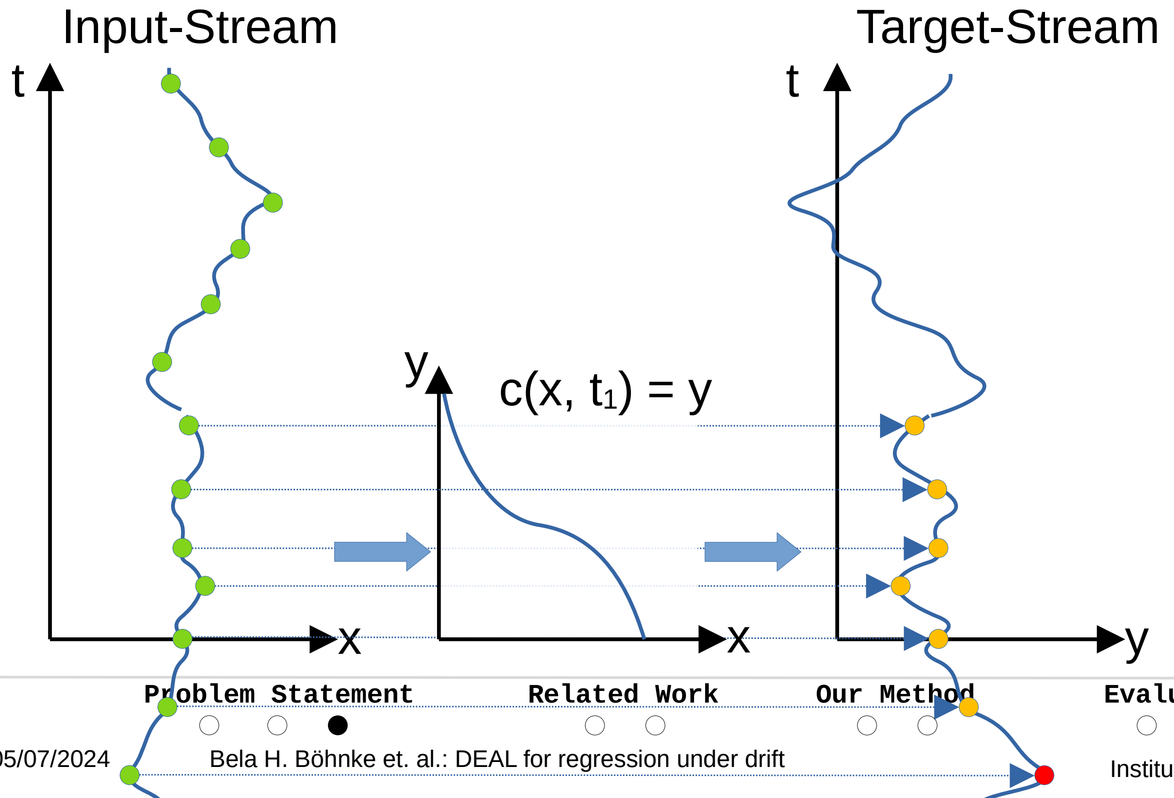
- Concept Drift



Problem Statement

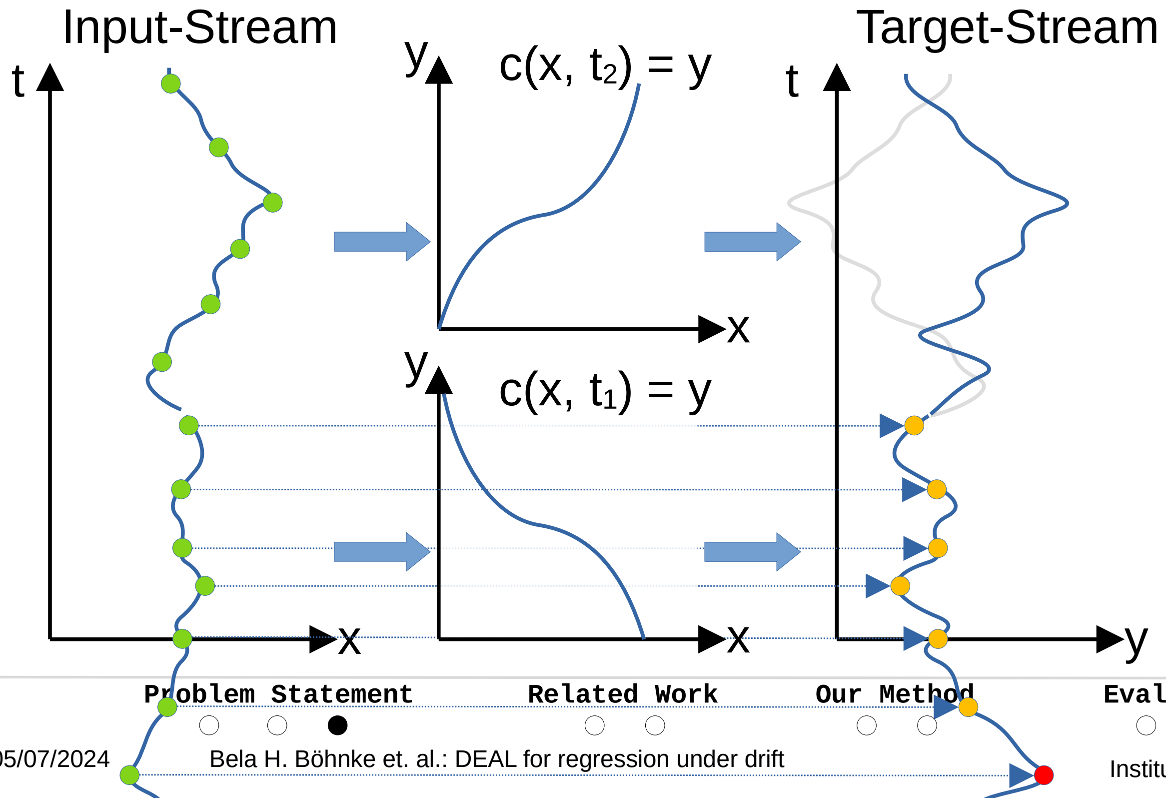
Karlsruhe Institute of Technology

- Concept Drift



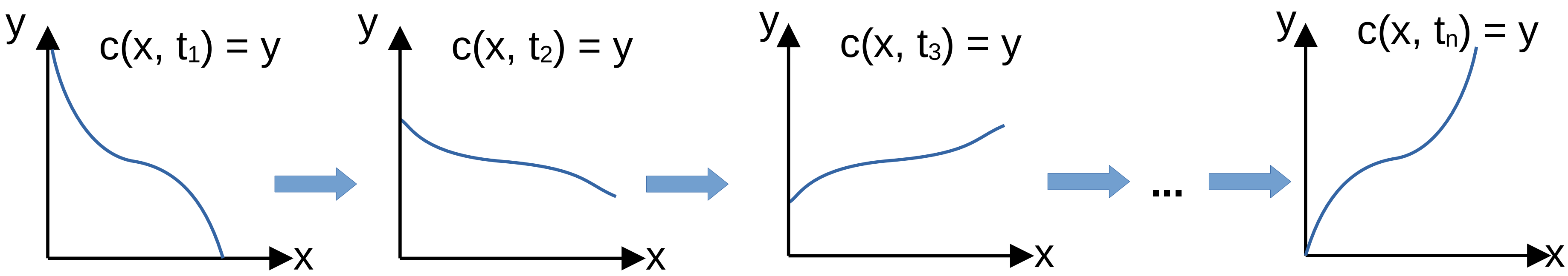
Problem Statement

- Concept Drift



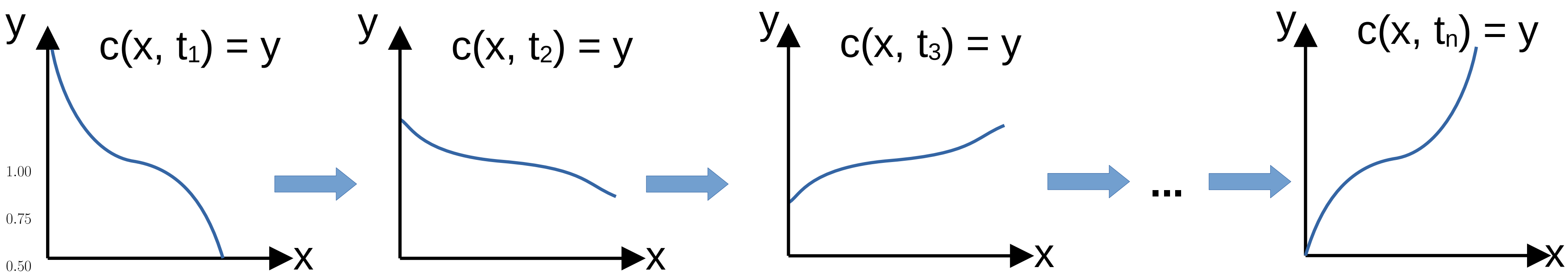
Problem Statement

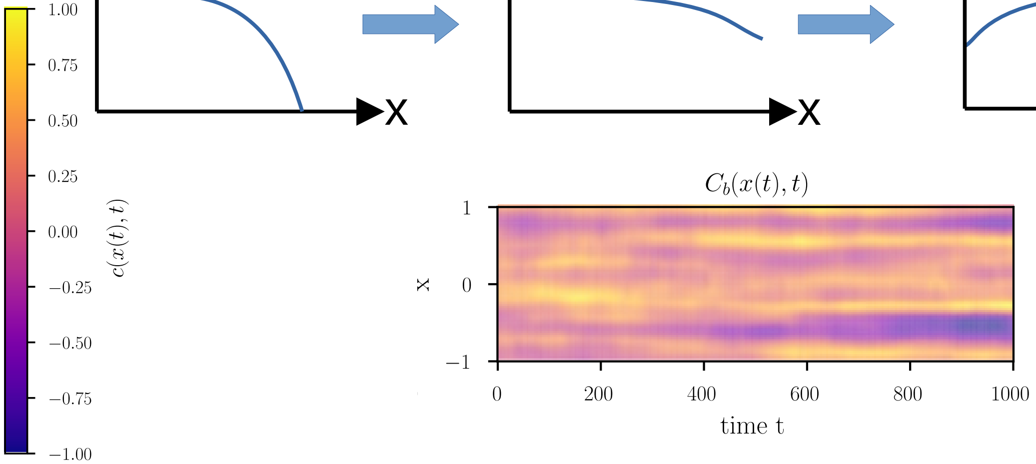
- Concept Drift



Problem Statement

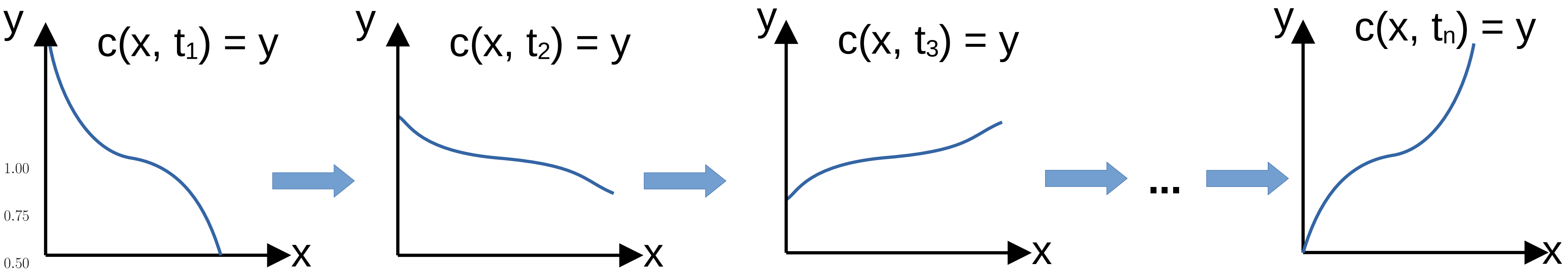
- Concept Drift

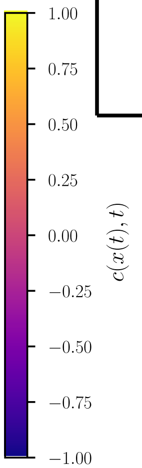




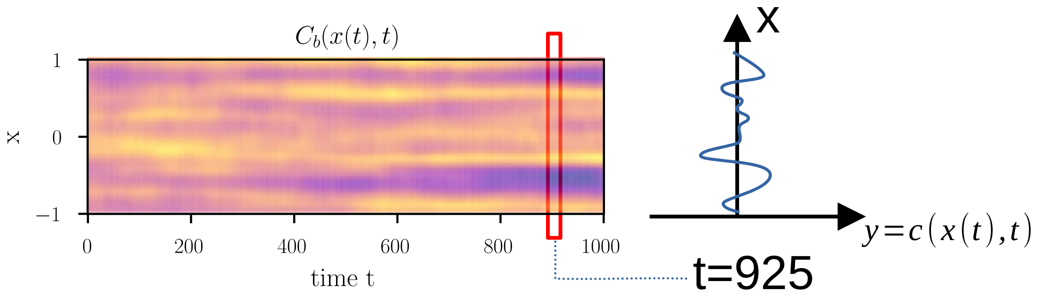
Problem Statement

- Concept Drift



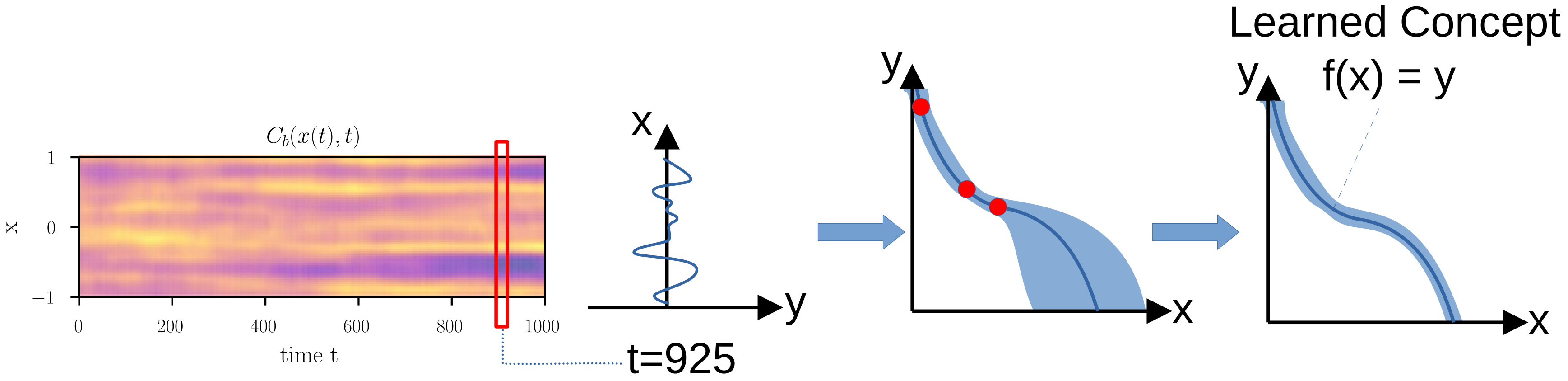


The value of y corresponds to the color.



Problem Statement

- The Challenge for Active Learning for Regression under Concept Drift:

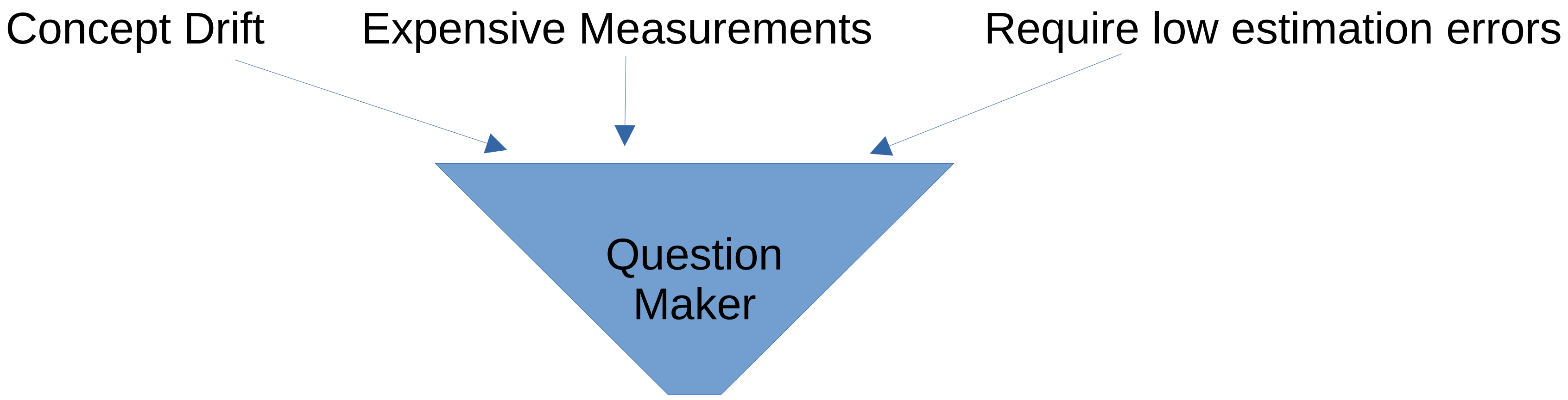


- Once uncertainty is low, method stops measuring and learning unaware of drift.
  → Method gets stuck.
- There is no way to detect a obsolete concept without expensive measurement of the target (such measurements are no longer performed).

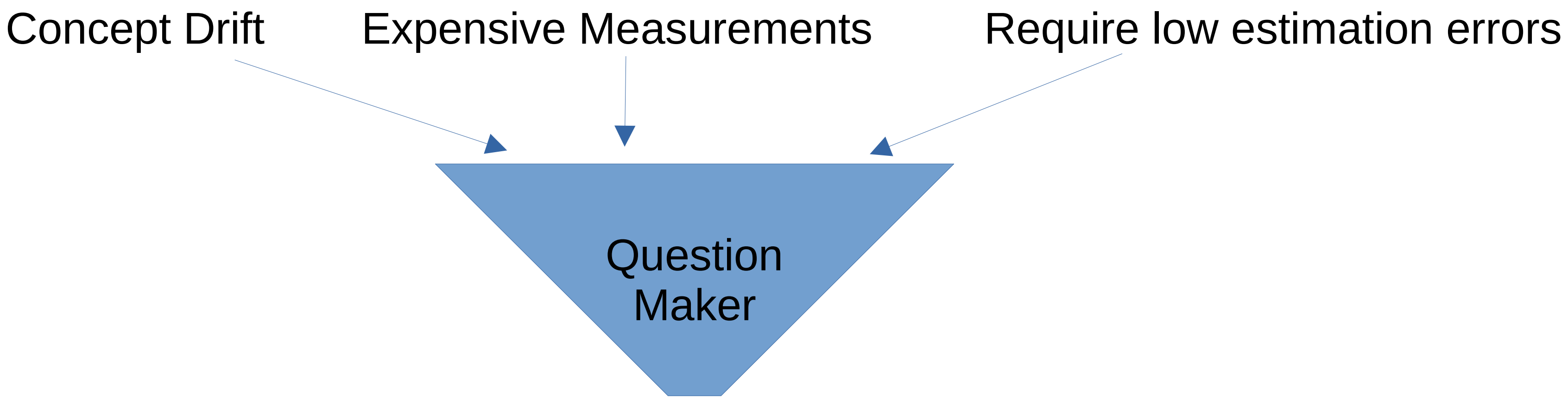
Problem Statement (summary)

- Concept Drift:
  - The learned function (concept) becomes obsolete after some time.
- Active Learning:
  - There is no way to detect a obsolete concept without measuring the target.
  - Measuring the target is expensive.
- User Requirement:
  - A user requires estimations for which estimation errors are below a “threshold of usefulness”
  - Estimations with higher error are harmful to follow-up tasks.
    (Think of crops planted in supposedly good soil that all die because of bad soil quality)

Research question

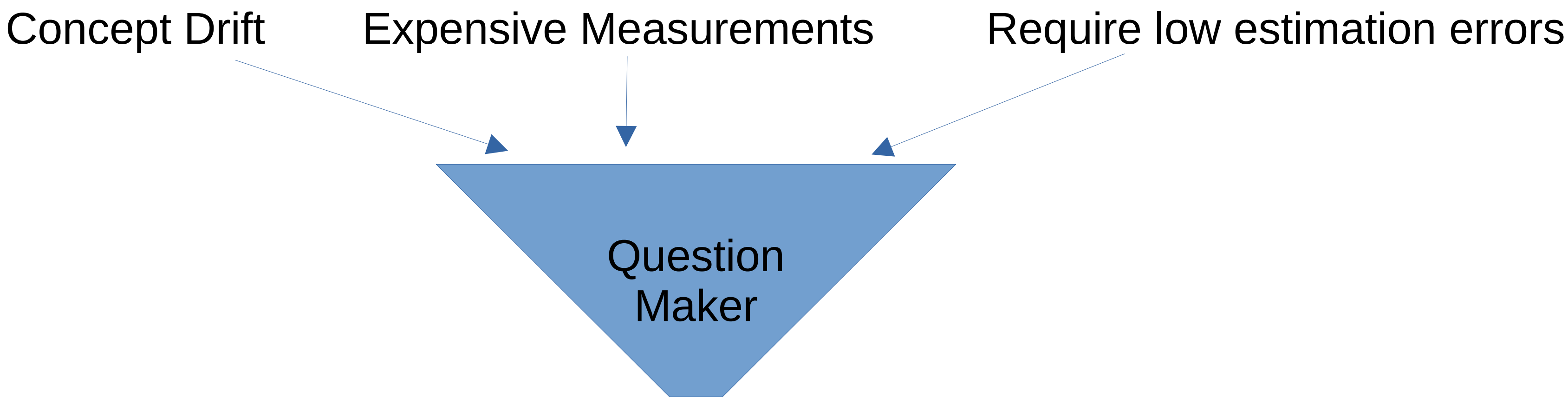


Research question



How can we decide how often to measure the target
(adapt the measurement frequency)
while:
Keeping the number of measurements to a minimum,
Keeping prediction errors below a user-required threshold?

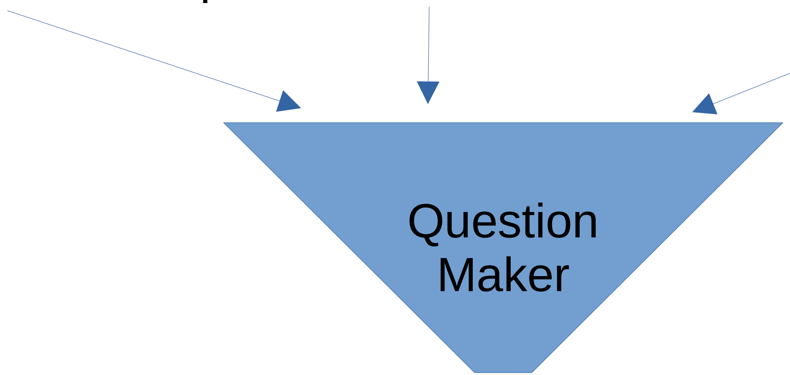
Research question



Adapt measurement frequency.
Minimize expensive measurements.
Satisfy user-required threshold.

Related Work

Concept Drift | Expensive Measurements | Require low estimation errors



Adapt measurement frequency.
Minimize expensive measurements.
Satisfy user-required threshold.

Related work answers some of these points separately, but has no answer for all in combination!

Related Work (summary)

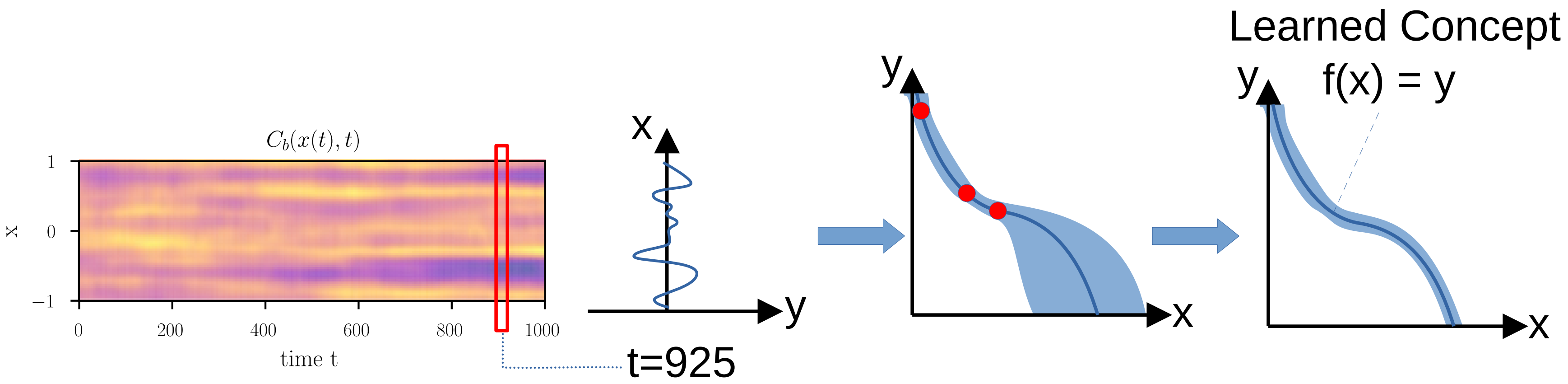
	Active learning for regression	Concept drift (no active learning)	Active learning for classification with drift	Methods in practice
Adapt measurement frequency:	✗	✗		✗
Minimize expensive measurements:		✗		✗
Satisfy user-required threshold:	✗		✗	✗

Our Method

	Our Method
Adapt measurement frequency:	✓
Minimize expensive measurements:	✓
Satisfy user-required threshold:	✓

Our Method

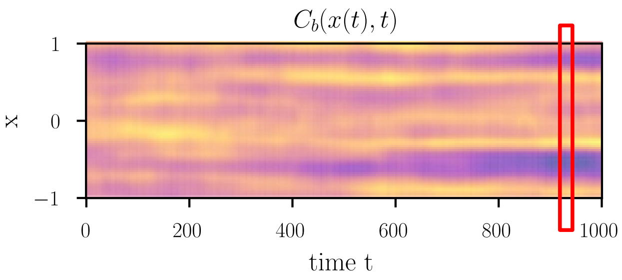
- Remember: Active Learning without considered drift:

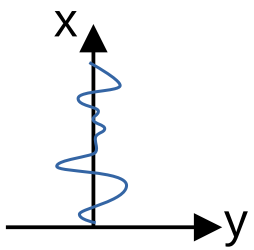


- Once uncertainty is low, method stops learning unaware of drift.
  → Method gets stuck.

Our Method

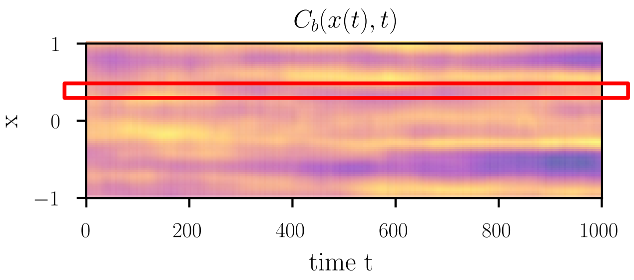
- We need to increase the uncertainty again! (preventing getting stuck)
- We learn statistics about the drift behavior:





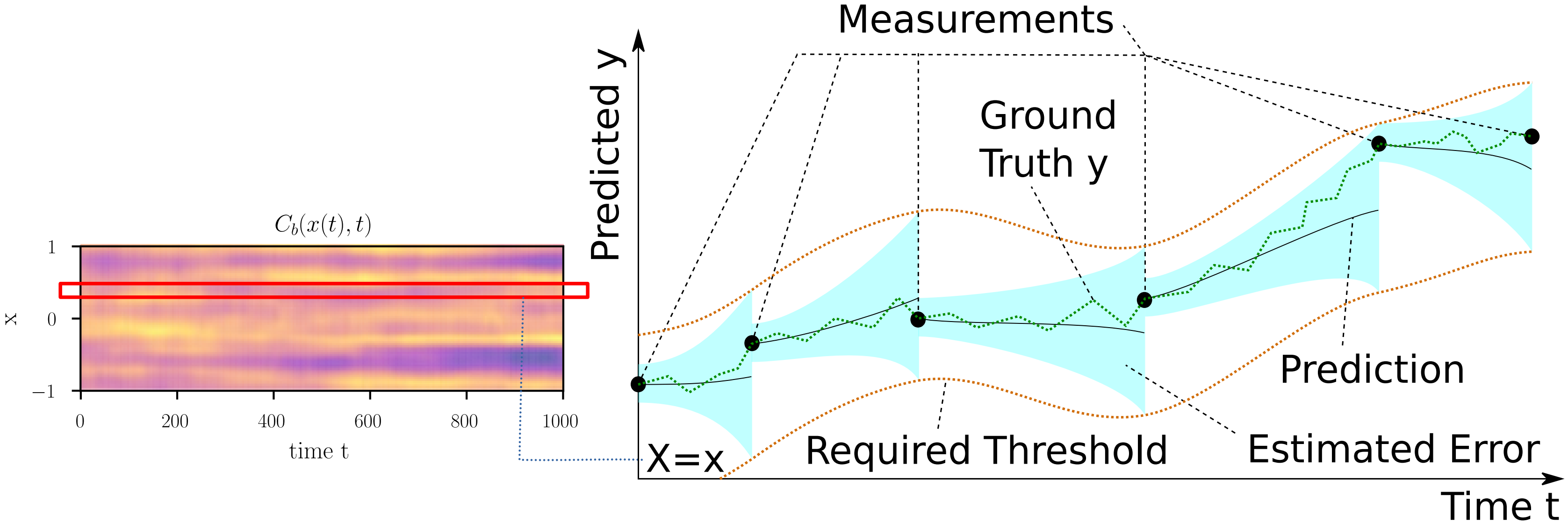
Our Method

- We need to increase the uncertainty again! (prevent getting stuck)
- We learn statistics about the drift behavior:



Our Method

- We need to increase the uncertainty again! (prevent getting stuck)
- We learn statistics about the drift behavior:



Our Method

- We use a N + 1 dimensional Gaussian Process Model.
- N [1 … 5] dimensions for the Input Stream X.
- 1 dimension for the time t.

- To model the increasing uncertainty over time we use a Brownian kernel $B(t)$.
- We use an RBF kernel $I(t)$ for modeling the input-target relation.
- We use an RBF kernel $W(t)$ for weighting / quantifying how much impact the drift has on the output y for a given input x.

$$C(x,t)=I(x)+W(x)B(t)$$

Our Method

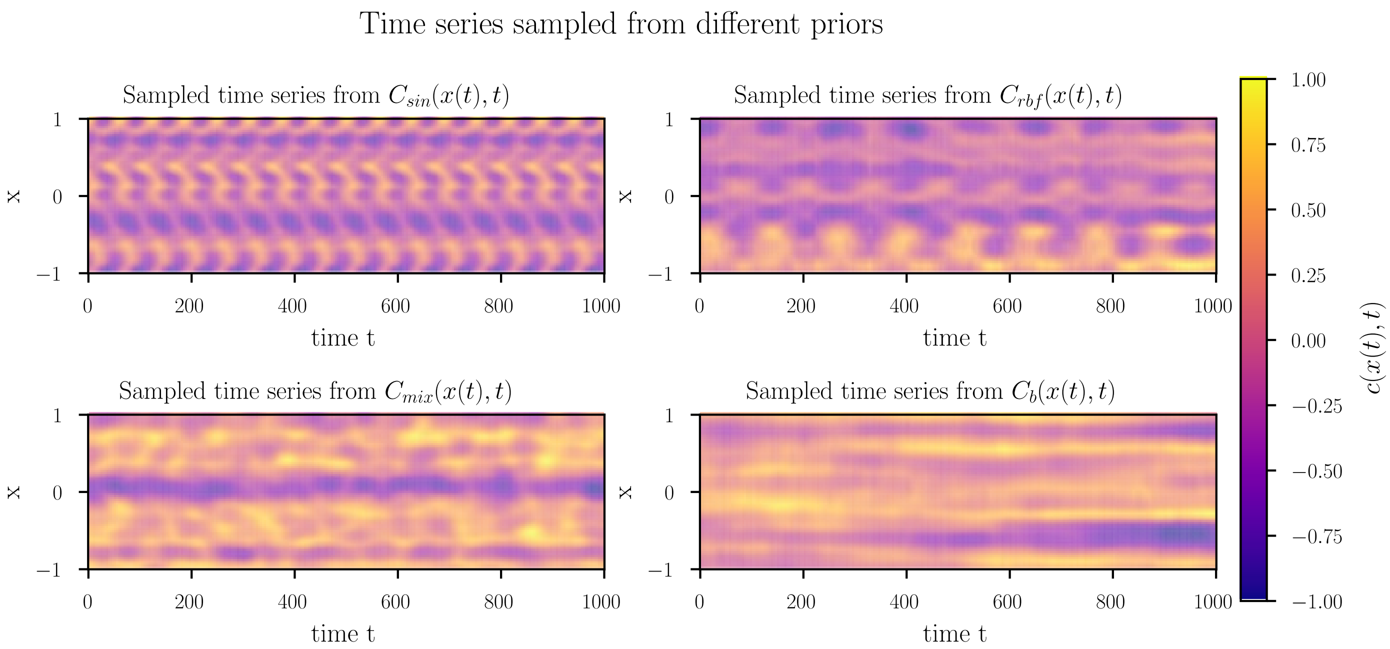
- Details:
  - We perform measurements once the estimated uncertainty reaches the user-required threshold.
  - We recalibrate the Gaussian Process Model using the measurement history and the new measurement.

Evaluation

- We evaluate on five datasets commonly used for classification and adapted them to regression.
- We evaluate on input streams with dimensions 1...5.
- We compare against four baseline:
  - Consecutive Measurement (as used in practice).
  - Classic Active Learning (not considering drift).
  - A change detection approach AAIL [24] (adapted from classification).
  - A second change detection approach considering detection errors.
- We evaluate each approach and each parameter configuration 50 times.

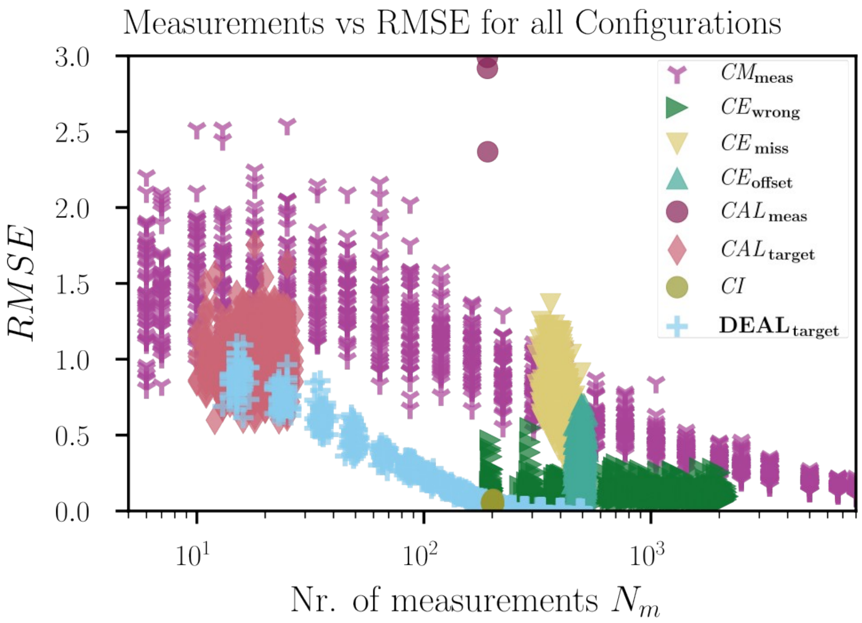
Evaluation

- Example one dimensional data:



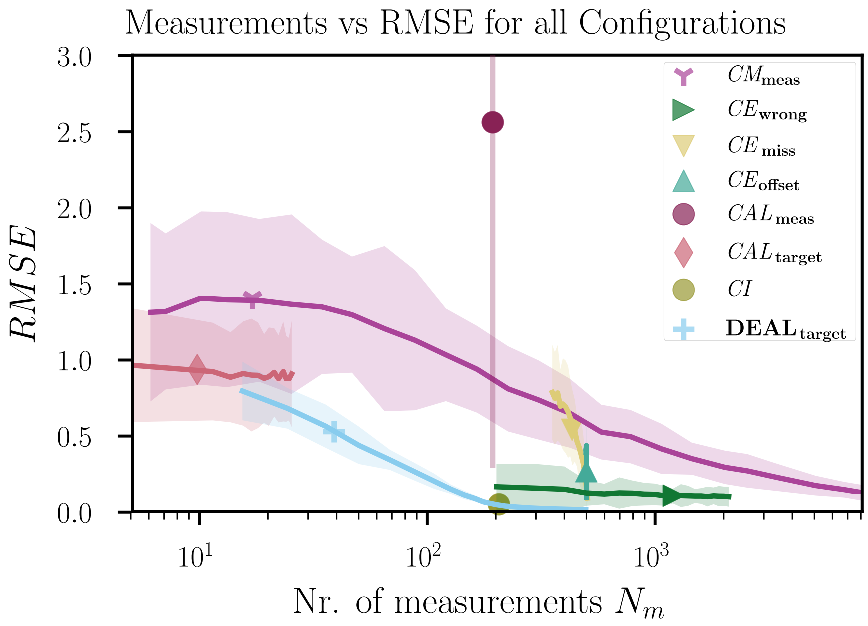
Evaluation

- Results



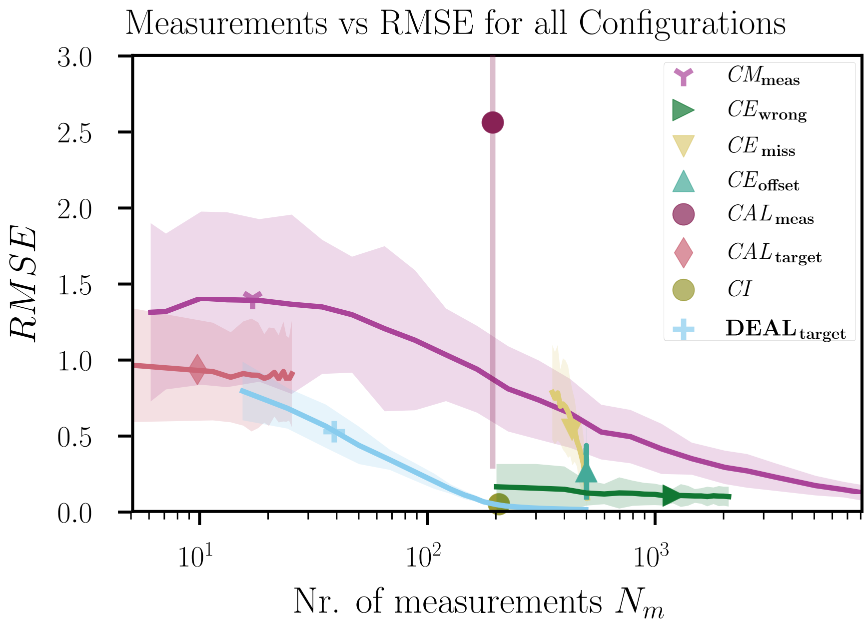
Evaluation

■ Results



Evaluation

Results

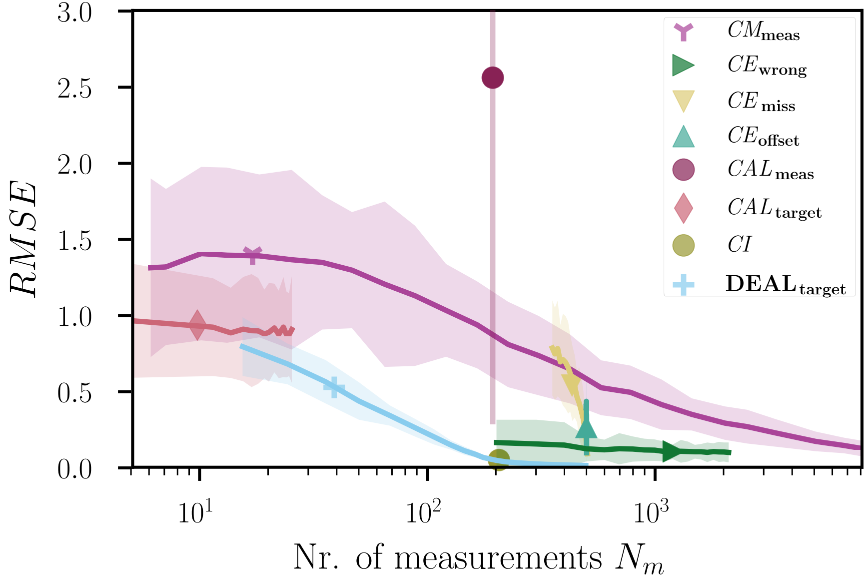


- Only DEAL and CM enable control of the user-required maximal error.

Evaluation

- Results

Measurements vs RMSE for all Configurations

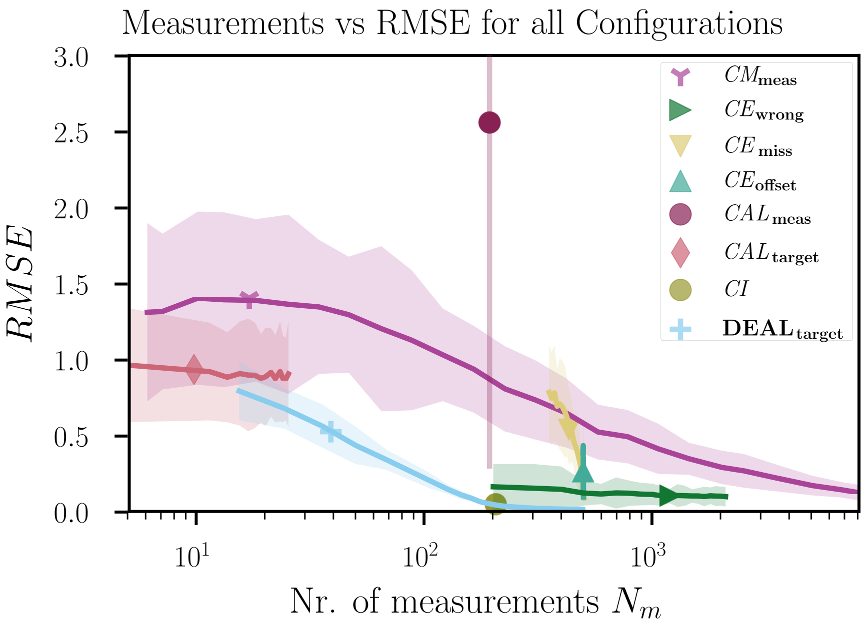


- Only DEAL and CM enable control of the user-required maximal error.
- Over all datasets DEAL has the most stable performance.

Evaluation

Karlsruhe Institute of Technology

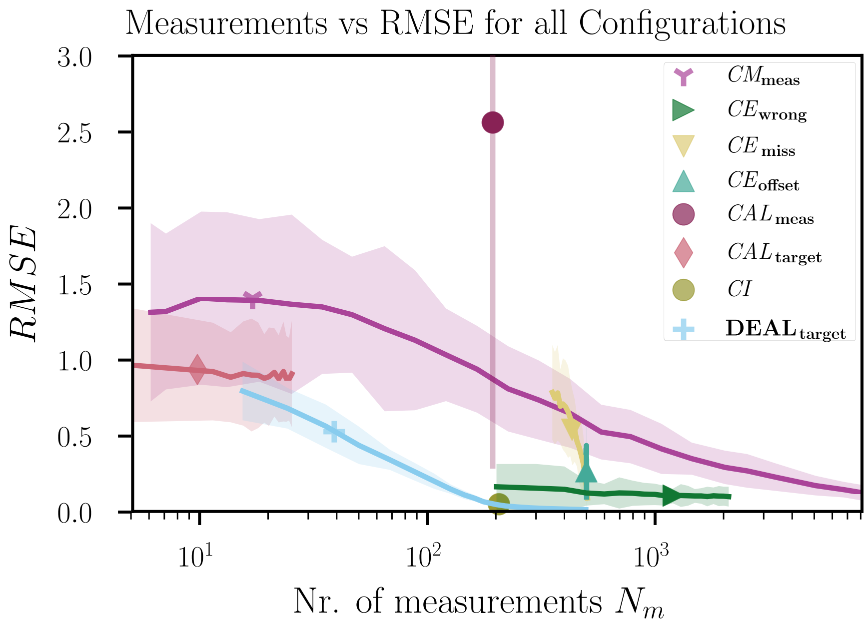
- Results



- Only DEAL and CM enable control of the user-required maximal error.
- Over all datasets DEAL has the most stable performance.
  - DEAL automatically adapts measurement frequency.

Evaluation

Results



- Only DEAL and CM enable control of the user-required maximal error.
- Over all datasets DEAL has the most stable performance.
  - DEAL automatically adapts measurement frequency.
- For a given error-threshold DEAL requires on average, 20 times fewer measurements.

Conclusion

- Challenges:
  - The relationship between input and target variables may drift due to environmental influences that are not observed.
  - Current work on active learning does not consider this for continuous variables.
- Our Contribution:
  - We proposed DEAL, a method that satisfies a given user-required prediction error threshold by adapting its measurement frequency to the drifting relationship.
  - DEAL requires, on average, 20 times fewer measurements than methods used in practice.
  - DEAL automatically adapts to changing drift behavior, preventing model degradation and improper set parameters.

References

1. Bachman, P., Sordoni, A., Trischler, A.: Learning algorithms for active learning. In: ICML (2017)

2. Bifet, A., Holmes, G., Kirkby, R., Pfahringer, B.: MOA: massive online analysis. J. Mach. Learn. Res. (2010)

3. C., K.F.: Introduction to stochastic calculus with applications. WSPC (2005)

4. Carne, G.D., Buticchi, G., Liserre, M., Vournas, C.: Load control using sensitivity identification by means of smart transformer. IEEE Trans. Smart Grid (2018)

5. Carne, G.D., Buticchi, G., Liserre, M., Vournas, C.: Real-time primary frequency regulation using load power control by smart transformers. IEEE Trans. Smart Grid (2019)

6. Dong, H., Jin, M., Ren-mu, H., Z.Y., D.: A real application of measurement-based load modeling in large-scale power grids and its validation. IEEE Trans. Power Systems (2009)

7. Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M., Bouchachia, A.: A survey on concept drift adaptation. ACM Comput. Surv. (2014)

8. GPy: A gaussian process framework in python. http://github.com/SheffieldML/GPy (since 2012)

9. Iwashita, A.S., Papa, J.P.: An overview on concept drift learning. IEEE Access (2019)

10. Iwata, T.: Active learning for regression with aggregated outputs. CoRR (2022)

11. João, G., Pedro, M., Gladys, C., Pedro, R.: Learning with drift detection. In: artif. intell. adv. – SBIA (2004)

12. Konyushkova, K., Sznitman, R., Fua, P.: Learning active learning from data. In: NIPS (2017)

13. Krawczyk, B., Cano, A.: Adaptive ensemble active learning for drifting data stream mining. In: IJCAI (2019)

14. Krawczyk, B., Pfahringer, B., Wozniak, M.: Combining active learning with concept drift detection for data stream mining. In: IEEE BigData (2018)

15. Kurlej, B., Wozniak, M.: Learning curve in concept drift while using active learning paradigm. In: ICAIS (2011)

16. Kurlej, B., Wozniak, M.: Active learning approach to concept drift problem. Log. J. IGPL (2012)

17. Lindgren, G., Rootzen, H., Sandsten, M.: Stationary Stochastic Processes for Scientists and Engineers. T&F (2013)

18. Liu, S., Xue, S., Wu, J., Zhou, C., Yang, J., Li, Z., Cao, J.: Online active learning for drifting data streams. IEEE Trans. Neural Networks Learn. Syst. (2023)

19. Liu, W., Zhang, H., Ding, Z., Liu, Q., Zhu, C.: A comprehensive active learning method for multiclass imbalanced data streams with concept drift. Knowl. Based Syst. (2021)

20. Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., Zhang, G.: Learning under concept drift: A review. IEEE Trans. Knowl. Data Eng. (2019)

References

21. Micchelli, C.A., Xu, Y., Zhang, H.: Universal kernels. J. Mach. Learn. Res. (2006)

22. Mohamad, S., Sayed Mouchaweh, M., Bouchachia, A.: Active learning for data streams under concept drift and concept evolution. In: ECML-PKDD (2016)

23. Montiel, J., Halford, M., Mastelini, S.M., Bolmier, G., Sourty, R., Vaysse, R., Zouitine, A., Gomes, H.M., Read, J., Abdessalem, T., et al.: River: machine learning for streaming data in python. JMLR (2021)

24. Park, C.H., Kang, Y.: An active learning method for data streams with concept drift. In: IEEE BigData (2016)

25. Riquelme, C., Johari, R., Zhang, B.: Online active linear regression via thresholding. In: AAAI (2017)

26. Settles, B.: Active Learning. Springer Cham (2012)

27. Shan, J., Zhang, H., Liu, W., Liu, Q.: Online active learning ensemble framework for drifted data streams. IEEE Trans. Neural Networks Learn. Syst. (2019)

28. Stefano, M., Bruno, S.: An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis. Struct. Saf. (2018)

29. Turab, L., V., B.P., Dezhen, X., Ruihao, Y.: Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. Npj Comput. Mater. (2019)

30. Viktor, L., Barbara, H., Wersing, H.: Knn classifier with self adjusting memory for heterogeneous concept drift. In: ICDM (2016)

31. Wu, D., Lin, C., Huang, J.: Active learning for regression using greedy sampling. Inf. Sci. (2019)

32. Yoo, D., Kweon, I.S.: Learning loss for active learning. In: CVPR (2019)

33. Zhang, H., Liu, W., Shan, J., Liu, Q.: Online active learning paired ensemble for concept drift and class imbalance. IEEE Access (2018)

34. Zliobaite, I.: Learning under concept drift: an overview. CoRR (2010)

35. Zliobaite, I., Bifet, A., Holmes, G., Pfahringer, B.: MOA concept drift active learning strategies for streaming data. In: WAPA. JMLR Proceedings (2011)

36. Zliobaite, I., Bifet, A., Pfahringer, B., Holmes, G.: Active learning with evolving streaming data. In: ECML/PKDD (3) (2011)

37. Zliobaite, I., Bifet, A., Pfahringer, B., Holmes, G.: Active learning with drifting streaming data. IEEE Trans. Neural Networks Learn. Syst. (2014)

Please checkout the accepted manuscript:
https://edouardfouche.com/publications/

Karlsruhe Institute of Technology

DEAL: Data-Efficient Active Learning for regression under drift

A First Method Aimed at the Gap of Regression Active Learning with Drift